

Change-point detection for a transient change and high-dimensional covering

Jack Noonan

2021

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy



School of Mathematics
Ysgol Mathemateg

*“To visualise 43-dimensional space I simply
visualise n -dimensional space and let n be 43”*

— Unknown

Acknowledgements

Firstly, I would like to thank my supervisors Anatoly and Andrey. Anatoly, you are the most impressive person I have ever met. It has been a privilege to work with you and I will be forever grateful for your support. I want to thank you for your patience, encouragement and your constant willingness to conduct research with me. Andrey, you helped me settle into my PhD studies and encouraged me early on when I needed it the most.

Secondly, I would like to thank my family. Mum and Dad, you have always been there when I needed you and I am so fortunate to have you. I would also like to thank my friends "The Fisherman". Throughout my undergraduate and postgraduate degrees, you have always provided the necessary distractions away from work.

Finally, I would also like to extend my gratitude to EPSRC for their financial support and to the School of Mathematics for providing me with everything I needed for my research.

Contents

Contents	ix
List of Figures	xv
List of Tables	xxiii
I Change-point detection	3
1 Online change-point detection for a transient change	5
1.1 Introduction	5
1.2 Permanent change in distributions	7
1.2.1 The CUSUM and Shiryaev-Roberts procedures	7
1.2.2 Evaluating ARL for CUSUM and SR tests	8
1.2.3 Optimality criteria	9
1.3 Transient change in distributions	11
1.3.1 A collection of procedures	11
1.3.2 Detecting a transient change in Gaussian random variables	13
1.3.3 Optimality criteria	21
1.3.4 Comparison of tests	25
1.3.5 An application to real world data	27
2 Approximations for the boundary crossing probabilities of moving sums of normal random variables	31
2.1 Introduction: Statement of the problem	31
2.2 Boundary crossing probabilities and related characteristics: discrete and continuous time	33
2.2.1 Standardisation of the moving sums	33
2.2.2 Correlation between $\xi_{n,L}$ and $\xi_{n+k,L}$	34
2.2.3 Continuous-time (diffusion) approximation	34
2.2.4 Diffusion approximations for the main characteristics of interest	35
2.2.5 Durbin and Poisson Clumping approximations for the BCP $P(T, h)$	36
2.3 Diffusion approximation with and without discrete-time correction; $M \leq L$	37
2.3.1 Diffusion approximation, formulation	37
2.3.2 Derivation of (2.3.2)	37

2.3.3	Discrete Time Correction	39
2.3.4	Simulation study, $T \leq 1$	41
2.4	Approximations for the BCP in continuous and discrete time; $M > L$	42
2.4.1	Glaz approximation for $\mathcal{P}_{\mathbb{X}}(M, h)$	42
2.4.2	Shepp's formulas	43
2.4.3	An alternative representation of the Shepp's formula (2.4.3) .	43
2.4.4	Joint density for the values $\{S(i)\}$ and associated transition densities	44
2.4.5	Correcting Shepp's formula (2.4.3) for discrete time	45
2.4.6	The Glaz-Shepp-Siegmund approximation	47
2.5	Approximations for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ through eigenvalues of integral operators; $M > L$	49
2.5.1	Continuous time: approximations for $F(T, h)$	49
2.5.2	Correcting approximation (2.5.4) for discrete time	50
2.6	Simulation study	52
2.6.1	Accuracy of approximations for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$	52
2.6.2	Approximation for the BCP in the case of non-normal moving sums	54
2.6.3	Approximation for the BCP in the case of moving weighted sums	55
2.7	Approximation using Extreme value theory; $M > L$	55
2.8	Approximating Average Run Length (ARL)	56
2.9	Appendix A: Proof of Lemma 2.2.1	58
2.10	Appendix B: Derivation of Durbin approximation	59
2.11	Appendix C: Derivation of (2.3.17)	59
3	Approximating Shepp's constants for the Slepian process	61
3.1	Introduction	61
3.2	Construction of approximations	62
3.2.1	Existence of Shepp's constants and the approximations derived from general principles	62
3.2.2	Shepp's formula for F_n	64
3.2.3	An alternative representation of the Shepp's formula (3.2.5) .	64
3.2.4	Approximating $\lambda(h)$ through eigenvalues of integral operators	65
3.2.5	Main approximations	67
3.2.6	Consistency of approximations when h is large	68
3.2.7	Numerical results	69
3.3	Comparison of the upper tail asymptotics for the Slepian process against some other stationary Gaussian processes	72
3.4	Appendix	73
3.4.1	Approximations for $\Lambda(h)$	73
3.4.2	An approximation for $F_2(h)$	74
4	First passage times for Slepian process with linear and piecewise linear barriers	77
4.1	Introduction	77

4.2	Linear barrier $a + bt$	79
4.2.1	An important auxiliary result	79
4.2.2	Linear barrier $a + bt$ with integer T	80
4.2.3	An alternative representation of formula (4.2.5)	80
4.2.4	Linear barrier $a + bt$ with non-integer T	81
4.3	Piecewise linear barrier with one change of slope	82
4.3.1	Boundary crossing probability	82
4.3.2	Two particular cases of Theorem 4.3.1	83
4.4	Piecewise linear barrier with two changes in slope	84
4.4.1	Boundary crossing probability	84
4.4.2	A particular case of Theorem 4.4.1	85
4.4.3	Another linear barrier with two changes in slope	86
4.5	Application to change-point detection	86
4.5.1	Formulation of the problem	86
4.5.2	Approximation for $\mathbb{E}_0(\tau(h))$	87
4.5.3	Approximating the power of the test	88
4.6	Appendix	91
4.6.1	Proof of Theorem 4.2.2	91
4.6.2	Proof of Theorem 4.3.1 (and Theorem 4.2.1)	93
4.6.3	Proof of Theorem 4.4.1	96
5	Power of the MOSUM test for online detection of a transient change in mean	99
5.1	Introduction: Statement of the problem	99
5.2	Formulation of MOSUM statistic and its characteristics	101
5.2.1	Definition of MOSUM test	101
5.2.2	Main characteristics of interest	101
5.2.3	Standardisation of MOSUM statistics	102
5.3	Approximation for ARL, the average run length to false alarm	102
5.4	Power of the test	103
5.4.1	Equivalent definition for the power of the test $\mathcal{P}_\xi(h, A, L)$	103
5.4.2	Continuous-time (diffusion) approximation for $\mathcal{P}_\xi(h, A, L)$	104
5.4.3	Case (a): $\lambda = 1$	106
5.4.4	Case (b): $\lambda > 1$	109
5.4.5	Case (c): $\lambda < 1$	114
5.5	Comparison of MOSUM and CUSUM tests	116
5.6	Robustness to non-normality and non-uniform weights	117
5.6.1	Robustness to non-normality	118
5.6.2	Robustness to non-uniform weights	119
5.7	Appendix A	119
6	Application to the Singular Spectrum Analysis (SSA) change-point detection algorithm	123
6.1	Introduction: Statement of the problem	123
6.2	Boundary crossing probabilities: discrete and continuous time	124
6.2.1	Reformulation of the problem	124

6.2.2	Motivation for the problem	125
6.2.3	Continuous time approximation	125
6.2.4	Correlation between $\mathbb{S}_{n;L,Q}$ and $\mathbb{S}_{n+1;L,Q}$	126
6.3	Approximations of the boundary crossing probabilities	127
6.3.1	Case of large Q and large L	127
6.3.2	Case of small Q and large L	129
6.4	Simulation study	130

II High-dimensional Covering 135

7 Covering of high-dimensional cubes and quantization 137

7.1	Introduction	137
7.1.1	Main notation	138
7.1.2	Main problem of interest	138
7.1.3	Two contradictory guiding principles and a compromise	139
7.1.4	Covering a cube by smaller cubes and quantization	140
7.1.5	Structure of the chapter and main results	141
7.2	Volume of intersection of a cube and a ball	141
7.2.1	The main quantity of interest	141
7.2.2	A generalization of the quantity (7.2.1)	142
7.2.3	Normal approximation for the quantity (7.2.1)	142
7.2.4	Improved normal approximation	144
7.2.5	Simulation study	145
7.3	Covering a cube by n balls	147
7.3.1	The main covering scheme	148
7.3.2	Theoretical investigation of Scheme 1	149
7.3.3	Simulation study for assessing accuracy of approximations (7.3.7) and (7.3.8)	150
7.3.4	Other schemes	151
7.3.5	Numerical comparison of schemes	152
7.4	Covering a cube by cubes	155
7.4.1	Volume of intersection of two cubes	155
7.4.2	Proportion of a cube covered by smaller cubes with random centers	156
7.5	Quantization	157
7.6	Appendix A: Several facts about d -dimensional balls and cubes	160
7.6.1	Volume of the ball	160
7.6.2	Radius of the ball of unit volume	161
7.6.3	Almost all the volume is near the boundary	161
7.6.4	The area of volume concentration in a cube	161
7.6.5	Squared norm of a random point in a cube	162
7.6.6	Distance between two random points in a cube	162
7.6.7	Volume of the intersection of two balls of the same radius	163
7.6.8	A direct computation of $C_{d,Z,r}$	163
7.7	Appendix B: Important auxiliary results	164

8	Efficient quantization and weak covering of high dimensional cubes	167
8.1	Introduction	167
8.1.1	Main notation	167
8.1.2	Main problems of interest	168
8.1.3	Relation between quantization and weak coverage	168
8.1.4	Re-normalization of the quantization error	169
8.1.5	Renormalised versions and formulation of optimal design problems	169
8.1.6	Thickness of covering	170
8.1.7	The design of the main interest	170
8.1.8	Structure of the rest of the chapter and the main results	171
8.2	Quantization	172
8.2.1	Reformulation in terms of the Voronoi cells	172
8.2.2	Voronoi cells for $\mathbb{D}_{n,\delta}$	172
8.2.3	Explicit formulae for the quantization error	173
8.3	Closed-form expressions for the coverage area with $\mathbb{D}_{n,\delta}$ and approximations	175
8.3.1	Reduction to Voronoi cells	175
8.3.2	Intersection of $[-1, 1]^d$ with one ball	176
8.3.3	Expressing $C_d(\mathbb{D}_{n,\delta}, r)$ through $C_{d,Z,r}$	176
8.3.4	Approximation for $C_d(\mathbb{D}_{n,\delta}, r)$	177
8.3.5	Simple bounds for $C_d(\mathbb{D}_{n,\delta}, r)$	177
8.3.6	‘Do not try to cover the vertices’	177
8.4	Numerical studies	178
8.4.1	Quantization and weak covering comparisons	179
8.4.2	Accuracy of covering approximation and dependence on δ	181
8.4.3	Stochastic dominance	182
8.5	Appendix A: Proof of Theorem 8.3.1	182
8.6	Appendix B: Proof of Lemma 8.3.2	185
8.7	Appendix C: Proof of Theorem 8.3.2.	187
9	Non-lattice covering and quantization of high dimensional sets	191
9.1	Introduction	191
9.2	Weak covering	195
9.2.1	Comparison of designs from the view-point of weak covering	195
9.2.2	Reduction to the probability of covering a point by one ball	196
9.2.3	Designs of theoretical interest	196
9.3	Approximation of $C_d(\mathbb{Z}_n, r)$ for Design 1	197
9.3.1	Normal approximation for $P_{U,\delta,\alpha,r}$	197
9.3.2	Refined approximation for $P_{U,\delta,\alpha,r}$	199
9.3.3	Approximation for $C_d(\mathbb{Z}_n, r)$ for Design 1	200
9.4	Approximating $C_d(\mathbb{Z}_n, r)$ for Design 2a	201
9.4.1	Normal approximation for $P_{U,\delta,0,r}$	202
9.4.2	Refined approximation for $P_{U,\delta,0,r}$	202
9.4.3	Approximation for $C_d(\mathbb{Z}_n, r)$	204

9.5	Approximating $C_d(\mathbb{Z}_n, r)$ for Design 2b	204
9.5.1	Establishing a connection between sampling with and without replacement: general case	205
9.5.2	Approximation of $C_d(\mathbb{Z}_n, r)$ for Design 2b.	205
9.6	Numerical study	206
9.6.1	Assessing accuracy of approximations of $C_d(\mathbb{Z}_n, r)$ and studying their dependence on δ	206
9.6.2	Comparison across α	206
9.7	Quantization in a cube	210
9.7.1	Quantization error and its relation to weak covering	210
9.7.2	Quantization error for Design 1	210
9.7.3	Quantization error for Design 2a	211
9.7.4	Quantization error for Design 2b	211
9.7.5	Accuracy of approximations for quantization error and the δ -effect	211
9.8	Comparative numerical studies of covering properties for several designs	211
9.8.1	Covering comparisons	214
9.8.2	Quantization comparisons	215
9.9	Covering and quantization in the d -simplex	217
9.9.1	Characteristics of interest	217
9.9.2	Numerical investigation of the δ -effect for d -simplex	218
9.10	Appendix: An auxiliary lemma	220
10	Summary of this thesis	223
10.0.1	Summary of Part one and next steps	223
10.0.2	Summary of Part two and next steps	224
	Bibliography	225

List of Figures

0.1	Dependence between chapters within Part one	2
0.2	Dependence between chapters within Part two	2
1.1	Empirical probabilities of reaching the barrier H (dashed black) and approximation (1.3.20) (solid blue). Left: $A = 1$, $M/l_1 = 4$ with $l_0 = 25$ and $l_1 = 50$. Right: $A = 0.5$, $M/l_1 = 25$ with $l_0 = 10$ and $l_1 = 20$	19
1.2	Empirical probabilities of reaching the barrier H (dashed black) and corresponding versions of approximation (1.3.23) (solid red). Left: $A = 1$, $M/l_1 = 1$ with (a) $l_1 = 10$ and (b) $l_1 = 50$. Right: $A = 1$, $M/l_1 = 2$ with (a) $l_1 = 10$ and (b) $l_1 = 50$	19
1.3	Empirical probabilities of reaching the barrier H (dashed black) and corresponding versions of approximation (1.3.23) (solid red). Left: $A = 1$, $M/l_1 = 5$ with (a) $l_1 = 10$ and (b) $l_1 = 50$. Right: $A = 1$, $M/l_1 = 10$ with (a) $l_1 = 10$ and (b) $l_1 = 50$	20
1.4	$\mathbb{E}_\nu S_{n,L}$ as a function of n	23
1.5	Barrier $h - Q(t; \gamma, \kappa)$ for $\lambda = 1$	24
1.6	Barrier $B(t; h, 0, -\gamma, \gamma)$	24
1.7	Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (thick dashed black) and its Approximation 1 (solid red) for two different values of h . Left: $h = 3$. Right: $h = 4$	25
1.8	Left: Power of three tests with $A = 1$ and $l = 10$ and ARL= 500. Right: Power of three tests with $A = 0.5$ and $l = 20$ and ARL= 500.	27
1.9	Typical pressure data	28
1.10	Left: Behaviour of y_t . Right: MOSUM statistic with $L = 50$ and ARL= 5000.	28
1.11	Left: Behaviour of y_t . Right: MOSUM statistic with $L = 150$ and ARL= 5000.	29
2.1	Empirical probabilities of reaching the barrier h and four approximations. Left: $L = 5$, $M = 5$, $T = 1$. Right: $L = 10$, $M = 5$, $T = 1/2$	41
2.2	Empirical probabilities of reaching the barrier h and four approximations. Left: $L = 100$, $M = 100$, $T = 1$. Right: $L = 200$, $M = 100$, $T = 1/2$	42

2.3	Empirical probabilities of reaching the barrier h (dashed black) and corresponding versions of Approximation 6 (solid red). Left: $T = 1$ with (a) $L = M = 5$ and (b) $L = M = 100$. Right: $T = 2$ with (a) $L = 5, M = 10$ and (b) $L = 100, M = 200$	48
2.4	Empirical probabilities of reaching the barrier $h = 2$ as a function of L (dashed black), uncorrected diffusion approximation $P(T, 2)$ (dot-dashed blue) and corresponding version of $P_L(T, h)$, which is Approximation 2 (solid red). Left: $M = L$ ($T = 1$). Right: $M = 2L$ ($T = 2$).	48
2.5	Empirical probabilities of reaching the barrier h (dashed black), Approximation 4 (solid blue) and Approximation 7 (solid red). Left: $T = 1.5$ with $L = 10$. Right: $T = 0.2$ with $L = 10$	49
2.6	Empirical probabilities of reaching the barrier h (dashed black), Approximation 8 (dashed green) and Approximation 9 (solid red). Left: $T = 10$ with (a) $L = 5$ and (b) $L = 100$. Right: $T = 50$ with (a) $L = 5$ and (b) $L = 100$	52
2.7	Empirical probabilities of reaching the barrier h (dashed black), Approximation 8 (dashed green) and Approximation 9 (solid red). Left: $\varepsilon_i \sim \text{Uniform}[0, 1]$ and $T = 10$ with (a) $L = 20$ and (b) $L = 100$. Right: $\varepsilon_i \sim \text{Laplace}[0, 1]$ and $T = 10$ with (a) $L = 20$ and (b) $L = 100$	54
2.8	BCP for the weighted sums (grey) against Approximation 9 for the BCP for non-weighted moving sums (red dotted line). Left: case (i) with $L = 20, M = 200, T = 10$. Right: case (ii) with $L = 20, M = 200, T = 10$	56
2.9	Empirical probabilities of reaching the barrier h (dashed black) and Approximation 10 (solid red). Left: $T = 100$ with $L = 20$. Right: $T = 5$ with $L = 10$	57
3.1	Lower and upper bounds (3.2.1) (red dotted lines) for $\Lambda(h)$ (solid black line).	63
3.2	Approximations and their relative errors as functions of h	69
3.3	Empirical versions of $F_T(h)$ (dashed black) and Approximation 7 (solid red).	71
3.4	Comparison of the upper tail asymptotics for several Gaussian stationary process	74
4.1	Graphical depiction of a general boundary $B_{T,T'}(t; a, b, b')$ with $b < 0$ and $b' > 0$	83
4.2	Barrier $B_{1,1}(t; a, -b, b)$ with $b > 0$	84
4.3	Barrier $B_{1,1}(t; a, 0, -b')$ with $b' > 0$	84
4.4	Barrier $B(t; h, 0, -\mu, \mu)$ with $\mu > 0$	85
4.5	Graphical depiction of the boundary $Q_\nu(t; h, \mu)$	89
4.6	Ratio $\gamma(x, h, \mu)/\gamma(0, h, \mu)$ for $h = 3$ and $\mu = 3$	89
5.1	Means of r.v. ε_j under the alternative.	101

5.2	$\mathbb{E}_1 S_{n,L}$ as a function of n	104
5.3	$\mathbb{E}_1 S(t)$ as a function of t	105
5.4	Barrier $h-Q(t; \gamma, \kappa)$ for $\lambda=1$	106
5.5	Barrier $B(t; h, 0, -\gamma, \gamma)$	106
5.6	Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (thick dashed black) and its Approximation 1 (solid red) for two different values of h . Left: $h = 3$. Right: $h = 4$	108
5.7	Barrier $h-Q(t; \gamma, \kappa)$ for $\lambda>1$	109
5.8	Barrier $B(t; h, 0, -\gamma, 0, \gamma)$	109
5.9	The transition densities $q_h^{(0)}(s)$ (dashed black) , $q_h^{(1)}(s)$ (solid red) and $q_h^{(2)}(s)$ (dotted blue) for two different values of h . Left: $h = 2$. Right: $h = 3$	112
5.10	Barrier $B_1(t; h, -\gamma, 0, \gamma)$	113
5.11	Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and its approximations (solid red). Left: $h = 3, \lambda = 1.5$ with Approximation 3. Right: $h = 3, \lambda = 2$ with Approximation 2.	114
5.12	Barrier $h-Q(t; \gamma, \kappa)$ for $\lambda<1$	115
5.13	Barrier $B_2(t; h, -\gamma, 0, \gamma)$	115
5.14	Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and Approximation 4 (solid red) Left: $h = 3, \lambda = 0.5$. Right: $h = 3, \lambda = 0.75$	115
5.15	Left: Comparison of power between the MOSUM test (solid black) as a function of λ and the CUSUM test (dot-dashed blue) with $A = 1.25$ and $l = 10$. Right: Power of the MOSUM test against η for different λ with $A = 1.25$ and $l = 10$: $\lambda < 1$ (red-dotted) $\lambda > 1$ (solid black).	117
5.16	Left: Comparison of power between the MOSUM test (solid black) as a function of λ and the CUSUM test (dot-dashed blue) with $A = 0.5$ and $l = 50$. Right: Power of the MOSUM test against η for different λ with $A = 0.5$ and $l = 50$: $\lambda < 1$ (red-dotted) $\lambda > 1$ (solid black).	118
5.17	Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and its approximations (solid red). Left: $\varepsilon_i \sim \text{Laplace}(0, 1)$ with $L = 5$. Right: $\varepsilon_i \sim \text{Laplace}(0, 1)$ with $L = 20$	118
5.18	Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and its (solid red). Left: $\varepsilon_i \sim \text{Uniform}[0, 1]$ with $L = 5$. Right: $\varepsilon_i \sim \text{Uniform}[0, 1]$ with $L = 20$	119
5.19	Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ with weights (solid grey) and Approximation 1 (solid red). Left: case (i) with $L = 20$ and $h = 3$. Right: case (ii) with $L = 20$ and $h = 3$	120
6.1	The weight function $w_{L,Q}(\cdot)$, $1 \leq Q \leq L$	124
6.2	The BCP for the weighted sum of normal r.v. and its approximations: $L = 150, Q = 50, M = 1000$	131
6.3	The BCP for the weighted sum of normal r.v. and approximations: $L = 100, Q = 50, M = 1000$	132

6.4	The BCP for the weighted sum of normal r.v. and its approximations: $L = 100, Q = 100, M = 2000$	132
6.5	The BCP for the weighted sum of normal r.v. and its approximations: $L = 100, Q = 5, M = 2000$	133
6.6	The BCP for the weighted sum of normal random variables and the Durbin approximation: $L = 300, Q = 1, T = 10$ (top) and $L = 300, Q = 1, T = 1$ (bottom)	134
7.1	$d = 10, Z = 0, r \in [1, 2.5]$	146
7.2	$d = 50, Z = 0, r \in [3.2, 4.9]$	146
7.3	$d = 10, Z = 0, r \in [0.95, 1.25]$	146
7.4	$d = 50, Z = 0, r \in [3.2, 3.5]$	146
7.5	$d = 10, Z$ is a vertex of $\mathcal{C}_d, r \in [2, 5]$	146
7.6	$d = 50, Z$ is a vertex of $\mathcal{C}_d, r \in [6.5, 9.5]$	146
7.7	$d = 10, Z$ is a vertex of $\mathcal{C}_d, r \in [1.9, 2.5]$	147
7.8	$d = 50, Z$ is a vertex of $\mathcal{C}_d, r \in [6.5, 7]$	147
7.9	Z is at half-diagonal with $\ Z\ = \frac{1}{2}\sqrt{10}$	147
7.10	Z is at half-diagonal, $\ Z\ = \frac{1}{2}\sqrt{50}$	147
7.11	Z is at half-diagonal, $\ Z\ = \frac{1}{2}\sqrt{10}$	147
7.12	Z is at half-diagonal, $\ Z\ = \frac{1}{2}\sqrt{50}$	147
7.13	$d = 10, Z \in \mathcal{S}_{10}(0, 1.5), r \in [1, 3.5]$	148
7.14	$d = 50, Z \in \mathcal{S}_{50}(0, 1.75), r \in [3.5, 5.5]$	148
7.15	$d = 10, Z \in \mathcal{S}_{10}(0, 1.5), r \in [1, 1.4]$	148
7.16	$d = 50, Z \in \mathcal{S}_{50}(0, 1.75), r \in [3.5, 3.75]$	148
7.17	$C_d(\mathbb{Z}_n, r)$ and approximations: $n = 128, r = 1.520$	150
7.18	$C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 1.291$	150
7.19	$C_d(\mathbb{Z}_n, r)$ and approximations: $n = 128, r = 2.460$	151
7.20	$C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 2.290$	151
7.21	$C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 4.020$	151
7.22	$C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 5.175$	151
7.23	Scheme 1: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$	153
7.24	Scheme 1: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$	153
7.25	Scheme 2: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$	154
7.26	Scheme 2: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$	154
7.27	Scheme 3: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$	154
7.28	Scheme 3: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$	154
7.29	Scheme 7: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$	155
7.30	Scheme 7: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$	155
7.31	$n = 50, r \in [0.7, 0.85]$ increasing by 0.05.	157
7.32	$n = 128, r \in [0.6, 0.8]$ increasing by 0.05.	157
7.33	$\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 128$	158
7.34	$\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 512$	158
7.35	$\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 128$	158
7.36	$\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 512$	158
7.37	$\mathbb{E}\theta_n$ with $n = 128$	160

7.38	$\mathbb{E}\theta_n$ with $n = 512$	160
7.39	$\mathbb{E}\theta_n$ with $n = 128$	160
7.40	$\mathbb{E}\theta_n$ with $n = 512$	160
8.1	$Q_d(\mathbb{D}_{n,\delta^*})$ and $Q_d(\mathbb{D}_{n,1/2})$ as functions of d and $Q_d(\mathbb{D}_n^{(0)}) = 1/12$; $d = 3, \dots, 50$	175
8.2	$Q_d(\mathbb{D}_{n,\delta})$ as a function of δ and $Q_d(\mathbb{D}_n^{(0)}) = 1/12$ with $d = 10$	175
8.3	$C_d(\mathbb{D}_{n,\delta}, r)$ with upper and lower bounds: $d = 20, \delta = 1/2$	178
8.4	$C_d(\mathbb{D}_{n,\delta}, r)$ with upper and lower bounds: $d = 100, \delta = 1/2$	178
8.5	$C_d(\mathbb{D}_{n,\delta}, r)$ with $R_{0.999}$ and R_1 : $d = 5$	179
8.6	$C_d(\mathbb{D}_{n,\delta}, r)$ with $R_{0.999}$ and R_1 : $d = 50$	179
8.7	$C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 5$	179
8.8	$C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 10$	179
8.9	$C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 50$	179
8.10	$C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 500$	179
8.11	$C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 3, r$ from 0.6 to 1 increasing by 0.1	181
8.12	$C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 5, r$ from 0.7 to 1.1 increasing by 0.1	181
8.13	$C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 7, r$ from 0.8 to 1.1 increasing by 0.075	181
8.14	$C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 10, r$ from 0.95 to 1.15 increasing by 0.05	181
8.15	$C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 15, r$ from 1.15 to 1.35 increasing by 0.05	182
8.16	$C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 50, r$ from 2.05 to 2.35 increasing by 0.075	182
8.17	Densities $f_d(R)$ for the design \mathbb{D}_{n,δ^*} ; $d = 5, 10, 20$	182
8.18	c.d.f.'s of R for the design \mathbb{D}_{n,δ^*} ; $d = 5, 10, 20$	182
8.19	$d = 5$: design $\mathbb{D}_n^{(0)}$ seems to stochastically dominate \mathbb{D}_{n,δ^*}	183
8.20	$d = 10$: design \mathbb{D}_{n,δ^*} stochastically dominates design $\mathbb{D}_n^{(0)}$	183
8.21	$d = 5$: design $\mathbb{D}_n^{(0)}$ stochastically dominates design \mathbb{S}_n	183
8.22	$d = 10$: design \mathbb{D}_{n,δ^*} stochastically dominates design \mathbb{S}_n	183
9.1	$C_d(\mathbb{Z}_n, r)$ with $r_{0.999}$ and r_1 : $d = 10, \mathbb{Z}_n$ is a 2^{d-1} -factorial design with $n = 2^{d-1}$	193
9.2	$C_d(\mathbb{Z}'_m, r)$ with $r_{0.999}$ and r_1 : $d = 10, \mathbb{Z}_m$ is a 2^d -factorial design	193
9.3	$P_{U,\delta,\alpha,r}$ and approximations: $d = 10, \alpha = 0.5$	200
9.4	$P_{U,\delta,\alpha,r}$ and approximations: $d = 20, \alpha = 0.5$	200
9.5	$P_{U,\delta,\alpha,r}$ and approximations: $d = 10, \alpha = 0.5$	200
9.6	$P_{U,\delta,\alpha,r}$ and approximations: $d = 10, \alpha = 1$	200
9.7	$P_{U,\delta,\alpha,r}$ and approximations: $d = 20, \alpha = 0.5$	201
9.8	$P_{U,\delta,\alpha,r}$ and approximations: $d = 20, \alpha = 1$	201
9.9	$P_{U,\delta,0,r}$ and approximations: $d = 10, seed = 10$	203
9.10	$P_{U,\delta,0,r}$ and approximations: $d = 20, seed = 10$	203

9.11	$P_{U,\delta,0,r}$ and approximations: $d = 10$, $seed = 10$	203
9.12	$P_{U,\delta,0,r}$ and approximations: $d = 10$, $seed = 15$	203
9.13	$P_{U,\delta,0,r}$ and approximations: $d = 20$, $seed = 10$	204
9.14	$P_{U,\delta,0,r}$ and approximations: $d = 20$, $seed = 15$	204
9.15	Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 10$, $\alpha = 0.5$, $n = 128$. .	207
9.16	Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20$, $\alpha = 0.1$, $n = 128$. .	207
9.17	Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20$, $\alpha = 0.5$, $n = 512$. .	207
9.18	Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20$, $\alpha = 0.1$, $n = 512$. .	207
9.19	Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 50$, $\alpha = 0.5$, $n = 512$. .	207
9.20	Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 50$, $\alpha = 0.1$, $n = 512$. .	207
9.21	Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 10$, $\alpha = 0$, $n = 128$. .	208
9.22	Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20$, $\alpha = 0$, $n = 128$. .	208
9.23	Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20$, $\alpha = 0$, $n = 512$. .	208
9.24	Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 50$, $\alpha = 0$, $n = 512$. .	208
9.25	Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 10$, $n = 128$. .	208
9.26	Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 10$, $n = 256$. .	208
9.27	Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 20$, $n = 512$. .	209
9.28	Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 20$, $n = 2048$. .	209
9.29	$d = 10$, $n = 512$, $r = 1.228$	209
9.30	$d = 10$, $n = 1024$, $r = 1.13$	209
9.31	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 1$, $n = 500$	212
9.32	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 0.5$, $n = 500$	212
9.33	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 0.5$, $n = 1000$	212
9.34	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 1$, $n = 1000$	212
9.35	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.4): $d = 50$, $\alpha = 0.1$, $n = 1000$	212
9.36	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.4): $d = 50$, $\alpha = 1$, $n = 1000$	212
9.37	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 10$, $\alpha = 0$, $n = 100$	213
9.38	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 10$, $\alpha = 0$, $n = 500$	213
9.39	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 20$, $\alpha = 0$, $n = 500$	213
9.40	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 50$, $\alpha = 0$, $n = 500$	213
9.41	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 10$, $n = 100$	213
9.42	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 10$, $n = 500$	213
9.43	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 20$, $n = 500$	214
9.44	$\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 20$, $n = 1000$	214
9.45	$C_d(\mathbb{Z}_n, r)$ as a function of r for several designs: $d = 10$, $n = 512$. .	215
9.46	$C_d(\mathbb{Z}_n, r)$ as a function of r for several designs: $d = 20$, $n = 1024$.	215
9.47	$d = 10$, $n = 512$: Design 2a with $\delta = 0.5$ stochastically dominates Design 3 with $\delta = 0.8$	216
9.48	$d = 10$, $n = 1024$: Design 2b with $\delta = 0.5$ stochastically dominates Design 3 with $\delta = 0.82$	216
9.49	\mathcal{S}_d and $\mathcal{S}_{d,1}^{(\delta)}$ with $d = 2$ and $\delta = 0.5$	218
9.50	\mathcal{S}_d and $\mathcal{S}_{d,2}^{(\delta)}$ with $d = 2$ and $\delta = 0.5$	218
9.51	$C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 5$, $n = 128$, r from 0.11 to 0.17 increas- ing by 0.02.	219

9.52	$C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 10, n = 512, r$ from 0.13 to 0.19 increasing by 0.02.	219
9.53	$C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 20, n = 1024, r$ from 0.13 to 0.17 increasing by 0.01.	219
9.54	$C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 50, n = 1024, r$ from 0.12 to 0.15 increasing by 0.01.	219
9.55	$C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 5, n = 128, r$ from 0.11 to 0.17 increasing by 0.02.	219
9.56	$C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 10, n = 512, r$ from 0.13 to 0.19 increasing by 0.02.	219
9.57	$C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 20, n = 1024, r$ from 0.13 to 0.17 increasing by 0.01.	220
9.58	$C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 50, n = 1024, r$ from 0.11 to 0.14 increasing by 0.01.	220
9.59	$\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S1: $d = 20, n = 1024$	220
9.60	$\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S1: $d = 50, n = 1024$	220
9.61	$\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S2: $d = 20, n = 1024$	220
9.62	$\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S2: $d = 50, n = 1024$	220

List of Tables

1.1	Approximations for $\mathbb{E}_\infty \tau_V(H)$ with $A = 1$	9
1.2	Approximations for $\mathbb{E}_\infty \tau_\xi(h)$ with $L = 10$	16
1.3	Approximations for $\mathbb{E}_\infty \tau_\xi(h)$ with $L = 50$	16
1.4	Approximations for $\mathbb{E}_\infty \tau_Z(H)$ with $l_0 = 25, l_1 = 50, A = 1$	19
1.5	Approximations for $\mathbb{E}_\infty \tau_Z(H)$ with $l_0 = 1, l_1 = 10, A = 1$	19
2.1	Relative error of the CDA with respect to the empirical BCP (in percent)	42
2.2	Values of $\lambda_{L,1}(h)$, $\lambda_{L,2}(h)$ and $\mu_L(h)$ with $L = 20$ for different h	53
2.3	Average values from 100 evaluations of Approximation 5 for different h along with maximum and minimum with $L = 5$ and $T = 100$	53
2.4	Average values from 100 evaluations of Approximation 5 for different h along with maximum and minimum with $L = 20$ and $T = 100$	53
2.5	Average values from 100 evaluations of Approximation 5 for different h along with maximum and minimum with $L = 100$ and $T = 100$	53
2.6	Approximations for $\text{ARL}_h(\mathbb{X})$ and $SD(\tau_h(\mathbb{X}))$ with $L = 10$	58
2.7	Approximations for $\text{ARL}_h(\mathbb{X})$ and $SD(\tau_h(\mathbb{X}))$ with $L = 50$	58
3.1	$\lambda^{(i)}(h), i = 0, 1, \dots, 7$, for different h	70
3.2	Relative errors of $\lambda^{(i)}(h), i = 0, 1, \dots, 6$, against $\lambda^{(7)}(h)$	70
3.3	Values of $F_T(h)$ and Approximations 4 and 7 for different values of h and T	71
3.4	Summary of results for an experiment involving $N = 10^5$ runs of $S(t)$ in $[0, 10)$ with $h = 2$	71
3.5	Approximations for $\Lambda(h)$ with accurate decimal digits in bold.	74
3.6	Values of I and its approximation \hat{I} for different values of h	75
4.1	$\mathcal{P}(h, \mu), \gamma_2(0, h, \mu)$ and $\gamma_1(0, h, \mu)$ for different μ for three choices of ARL.	90
6.1	Threshold for a given BCP for the weighted sum of normal r.v. and approximations: $L = 150, Q = 50, M = 1000$	131
6.2	Threshold for a given BCP for the weighted sum of normal r.v. and its approximations: $L = 100, Q = 50, M = 1000$	131
6.3	Threshold for a given BCP for the weighted sum of normal r.v. and approximations: $L = 100, Q = 100, M = 2000$	133

6.4	Threshold for a given BCP for the weighted sum of normal r.v. and approximations: $L = 100, Q = 5, M = 2000$	133
6.5	Threshold for a given BCP for the weighted sum of normal r.v. and Durbin approximation: $L = 300, Q = 1, T = 1$	134
7.1	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 10$. .	153
7.2	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 20$. .	153
7.3	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 50$. .	154
7.4	Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across schemes and n for $d = 10$	159
7.5	Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across schemes and n for $d = 20$	159
7.6	Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across schemes and n for $d = 50$	159
7.7	Radius of the ball of unit volume for different dimensions	161
8.1	Normalised mean squared quantization error Q_d for four designs and different d	180
8.2	Normalised statistic $R_{1-\gamma}$ across d with $\gamma = 0.01$ (value in brackets corresponds to optimal δ)	180
9.1	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 5$. . .	209
9.2	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 10$. .	209
9.3	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 10$. .	214
9.4	Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 20$. .	214
9.5	Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across selected designs; $d = 10$	216
9.6	Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across selected designs; $d = 20$	216

Preface

This thesis is divided into two parts. Part one is the major contribution of this thesis and considers the topic of change-point detection. The majority of research in Part one focuses on deriving certain quantities for a particular change-point detection procedure. The derivation of these quantities relies on the ability to evaluate (or approximate) complicated boundary crossing probabilities for a particular Gaussian process. Part two is a topic of extra interest and studies the covering and quantization of high dimensional sets. As the main problem, we consider the covering and quantization of a d -dimensional cube by n balls with reasonably large d (10 or more) and reasonably small n , like $n = 100$. When considering covering problems, the full coverage is not enforced but instead, only 95% or 99% coverage is desired. The results of Part two establish that efficient covering schemes have several important properties which are not seen in small dimensions and in asymptotical considerations, for very large n . One of these properties can be termed ‘do not try to cover the vertices’ as the vertices of the cube and their close neighbourhoods are very hard to cover and for large d there are too many of them. The structure and content of this thesis is based on the eight published papers of the author; the relevant paper is referenced at the start of each chapter. Let us discuss the structure in more detail.

Part one is comprised of Chapters 1 – 6. Chapter 1 provides an introduction to the change-point problem considered in this thesis and offers a survey of the current state of the field. It also provides a survey of the main findings of the author and can be used to summarise Part one. Chapter 2 studies boundary crossing probabilities for a particular Gaussian process and a number of accurate approximations are derived. One of the key techniques is to generalise the sequential analysis results of David Siegmund. Chapter 3 studies constants related to this Gaussian process; their existence was posed by Larry Shepp. Chapter 4 derives previously unseen boundary-crossing probabilities which are necessary for evaluating the power of the change-point procedures in continuous and discrete time; Chapter 5 focuses solely on approximating power in discrete time. Chapter 6 discusses boundary-crossing probabilities related to the popular Singular Spectrum Analysis change-point detection procedure.

Part two is comprised of Chapters 7 – 9. Chapter 7 studies the covering and quantization properties of a number of schemes/designs. For a particular random design, approximations to these two quantities are offered. Chapter 8 focuses on a particular strongly performing set of points strongly related to the checkerboard lattice studied by John Conway and Neil Sloane; exact expressions for quantization and accurate approximations for its covering properties are derived. In Chapter 9, a

generalisation of results in Chapter 7 is presented.

In Figure 0.1, the dependence between the chapters of Part one is summarised. If an arrow points from Chapter A to Chapter B, then some important results of Chapter B are used within Chapter A. As previously mentioned, Chapter 1 acts as both a literature review and summarises the main findings from Part one. As a result, it is dependent on many later chapters. Although Chapter 6 is not directly related to any of the previous Chapters 1-5, the change-point procedure studied in Chapter 6 inspired the author's research in the area of change-point detection and has some (weak) connections to the one studied in Chapters 2-5.

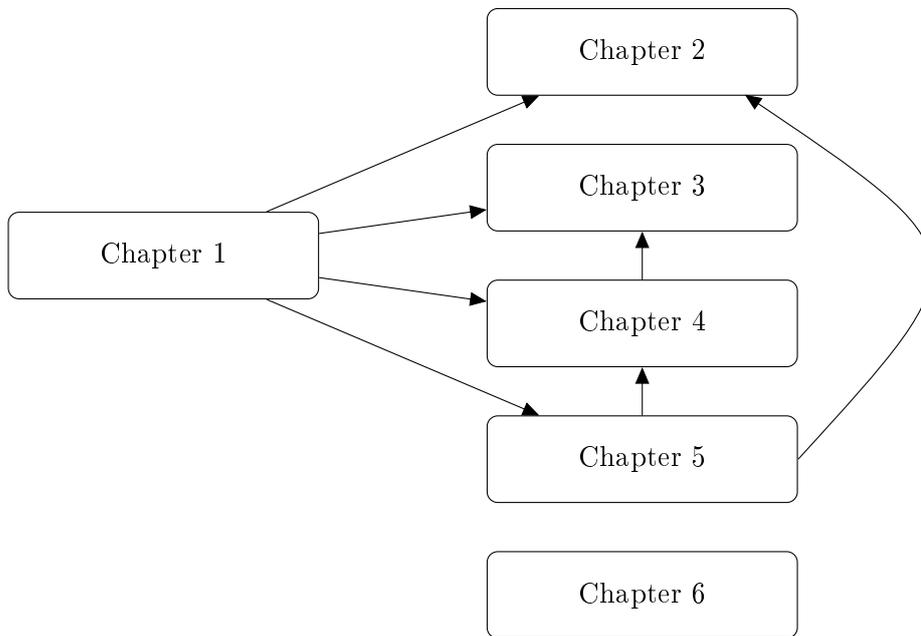


Figure 0.1: Dependence between chapters within Part one

In Figure 0.2, the dependence between the chapters of Part two is summarised. The numerical results of Chapter 7 inspired the research conducted in Chapters 8 and 9 and is the reason for the dependence.

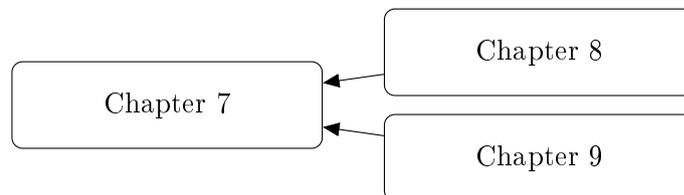


Figure 0.2: Dependence between chapters within Part two

To ensure a good standard of reproducibility, the R code used to support the findings of this thesis is available at <https://github.com/JackNoonan/Change-point> (or can be obtained from the author by an email request at Noonanj1@cardiff.ac.uk).

Part I

Change-point detection

Chapter 1

Online change-point detection for a transient change

Abstract

In this chapter, a popular online change-point problem of detecting a transient change in distributions of independent random variables is considered. This will be the main problem considered in Part one of this thesis. For this change-point problem, several change-point procedures are formulated and some advanced results for a particular procedure are surveyed. Some new approximations for the average run length to false alarm are offered and the power of these procedures for detecting a transient change in mean of a sequence of normal random variables are compared.

1.1 Introduction

The subject of change-point detection (or statistical quality control) is devoted to monitoring and detecting changes in the structure of a time series. This chapter and Part one of this thesis considers a popular online change-point problem of detecting a change in distribution of a sequence of independent random variables. Online change-point problems are concerned with monitoring the structure of a random process(es) whose observations arrive sequentially. For these problems, any good monitoring procedure should reliably alert the user to unexpected changes as soon as possible or with highest probability, subject to a tolerance on false alarms. In this chapter, we will assume the distributions before and after a change-point have explicit probability density functions. This is a common assumption in the field, see for example [131, 141].

Let y_1, y_2, \dots be a sequence of independent random variables arriving sequentially. The purpose of this chapter is to discuss tests for the hypothesis that y_i ($i = 1, 2, \dots$) are identically distributed with some probability density function (pdf) $f(y)$ against the alternative that at some unknown change point $0 \leq \nu < \infty$, the random variables y_1, y_2, \dots, y_ν and $y_{\nu+l+1}, y_{\nu+l+2}, \dots$ are identically distributed with density $f(y)$ and $y_{\nu+1}, y_{\nu+2}, \dots, y_{\nu+l}$ are identically distributed with pdf $g(y)$ such that $g(y) \neq f(y)$. Here, l is length of the change-point period (signal) and can be known or unknown.

Under a standard hypothesis testing framework, the null hypothesis is $\mathbb{H}_\infty: \nu = \infty$ and hence the pdf $f(y)$ is the density of y_i for all $i = 1, 2, \dots$

The alternative hypothesis is $\mathbb{H}_\nu: 0 \leq \nu < l \leq \infty$ and therefore

$$\mathbb{H}_\nu: \begin{cases} y_i \text{ have density } f(y) & \text{if } i \leq \nu \text{ or } i > \nu + l \\ y_i \text{ have density } g(y) & \text{if } \nu < i \leq \nu + l \end{cases}$$

with $i = 1, 2, \dots$. Under \mathbb{H}_ν , the arrival time of the signal is $\nu + 1$ (it is unknown). Most classical results assume f and g are known completely; by this, we mean no nuisance parameters are present in the distributions. In later sections, tests designed to tackle the change-point problem with nuisance parameters are discussed.

A thorough introduction to the field of online (quickest) change-point detection mainly for the case of $l = \infty$ can be found in, for example, [4, 102, 131, 143]. Some of the most popular online change-point algorithms used in practice are Shewhart's \bar{X} -chart [115], the CUSUM algorithm [87], the Shiryaev-Roberts procedure [109, 116] and the Exponentially Weighted Moving Average (EWMA) chart [110]. The case of $l = \infty$, and hence when a change in distribution occurs it does so permanently, is by far the most popular scenario considered in the change-point literature; a number of influential papers are [37, 60, 68, 75, 92, 108]. The CUSUM and Shiryaev Roberts procedures benefit with their simplicity and proven optimality under suitable optimality criteria; these two procedures will be the focus of discussion for the case $l = \infty$. The case of finite l , and hence when a change occurs it does so temporarily, has seen considerable attention in the past, see [5, 6, 18, 33, 34, 39, 40, 59]. More recently it has been the focus of attention in the papers of [83, 84, 145]. Examples of areas where detecting a transient change in distributions is extremely important can be found in radar and sonar [8, 101, 126], nondestructive testing [111], and medicine [10]. Non-parametric online change-point detection methods have also become very popular [15, 74]. For the state of the art techniques for multiple change-point detection, see [17, 25, 26, 56].

This chapter is organised as follows. In Section 1.2, we survey results for $l = \infty$ and discuss known optimality results for the CUSUM and Shiryaev-Roberts procedures. This section contains well known classical results but is included to introduce the reader to change-point concepts that will be used when considering the transient change-point problem. In Section 1.3, we assume $l < \infty$ and discuss a number of online tests for transient changes; the likelihood ratio test providing the inspiration behind all tests. In this section, we compare procedures when applied for detecting a temporary change in mean of a sequence of Gaussian random variables. We also apply tests for monitoring stability of components used in the Oil and Gas industry.

Throughout this chapter we shall use the notation \Pr_∞ and \mathbb{E}_∞ to denote probability and expectation under \mathbb{H}_∞ . Under the alternative \mathbb{H}_ν , we shall use the notation \Pr_ν and \mathbb{E}_ν to denote probability and expectation assuming the change-point occurs at $\nu < \infty$.

1.2 Permanent change in distributions

In this section, we assume $l = \infty$; if a change occurs, it does so permanently. Suppose y_1, y_2, \dots, y_n have been sampled. The likelihood ratio for testing \mathbb{H}_∞ against \mathbb{H}_ν is

$$\Lambda_{\nu,n} = \prod_{i=\nu+1}^n \frac{g(y_i)}{f(y_i)}$$

assuming $\nu < n$, otherwise $\Lambda_{\nu,n} = 1$.

1.2.1 The CUSUM and Shiryaev-Roberts procedures

By maximising the statistic $\Lambda_{\nu,n}$ over all possible locations of ν , we obtain the cumulative sum (CUSUM) statistic

$$V_n := \max_{0 \leq \nu \leq n-1} \Lambda_{\nu,n}, \quad n \geq 1. \quad (1.2.1)$$

The CUSUM stopping rule (when to alert the user to a potential change-point) is

$$\tau_V(H) := \inf\{n \geq 1 : V_n > H\}. \quad (1.2.2)$$

An appealing property of statistic (1.2.1) is the recursive property

$$V_n = \max\{V_{n-1}, 1\} \cdot \frac{g(y_n)}{f(y_n)}, \quad V_0 = 1.$$

The threshold H in $\tau_V(H)$ is chosen on the users tolerance to false alarm risk. Page [87] and Lorden [68] measured false alarm risk through the Average Run Length to false alarm (ARL). The ARL criterion corresponds to choosing H such that $\mathbb{E}_\infty \tau_V(H) = C$, where C is a pre-defined value chosen by the user but is typically large. How to compute $\mathbb{E}_\infty \tau_V(H)$ will be discussed later in this section.

The famous CUSUM chart of Page [87] introduces a reflective barrier at zero. This procedure is defined as:

$$P_n := \max\left\{P_{n-1} + \log \frac{g(y_n)}{f(y_n)}, 0\right\}, \quad P_0 = 0. \quad (1.2.3)$$

The statistics (1.2.3) and $\log V_n$ are equivalent on the positive half plane and hence the rule

$$\tau_P(\log(H)) = \inf\{n \geq 1 : P_n > \log H\},$$

and $\tau_V(H)$ are equivalent for $H > 1$. The stopping rule τ_V is more general than τ_P as thresholds $H \leq 1$ are permissible. An approximation for $\mathbb{E}_\infty \tau_P(H)$ for general distributions f and g was derived in [100]. Let $I_f := -\mathbb{E}_\infty(\log(g(y_1)/f(y_1)))$ and $I_g = \mathbb{E}_0(\log(g(y_1)/f(y_1)))$ (to compute I_g we assume the change-point occurs at time zero). Then

$$\mathbb{E}_\infty \tau_P(H) \simeq \frac{e^H}{I_g \zeta^2} - \frac{H}{I_f} - \frac{1}{I_g \zeta}. \quad (1.2.4)$$

Here the constant ζ is called the limiting exponential overshoot. Let $Z_n = \sum_{i=1}^n \log(g(y_i)/f(y_i))$ be the random walk. For a non-negative barrier a , define the stopping rule $\tau_a := \inf\{n \geq 1 : Z_n > a\}$ and define the excess over the barrier by $\kappa_a := Z_{\tau_a} - a$. Then $\zeta := \lim_{a \rightarrow \infty} \mathbb{E}_0[e^{-\kappa_a}]$. It was shown in [117, Ch. VIII] that

$$\zeta = \frac{1}{I_g} \exp \left\{ - \sum_{k=1}^{\infty} \frac{1}{k} [\Pr_{\infty}(Z_k > 0) + \Pr_0(Z_k \leq 0)] \right\}.$$

The approximation (1.2.4) seems extremely accurate. For example, suppose pre-change observations are i.i.d. $N(0, 1)$ random variables and post-change observations are i.i.d. $N(A, 1)$ for some known $A > 0$. We have

$$f(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2), \quad g(y) = \frac{1}{\sqrt{2\pi}} \exp(-(y-A)^2/2). \quad (1.2.5)$$

For $A = 1$, Monte Carlo simulations with 100,000 iterations provide $\mathbb{E}_{\infty} \tau_P(4.39) = 500$. Application of the approximation in (1.2.4) provides 498. The draw back of the approximation in (1.2.4) is that ζ requires expensive numerical evaluation.

To construct the Shiryaev-Roberts (SR) procedure, define the generalised Bayesian detection statistic as:

$$R_n := \sum_{\nu=0}^n \Lambda_{\nu, n}. \quad (1.2.6)$$

Then the SR test is:

$$\tau_R(H) := \inf\{n \geq 1 : R_n > H\},$$

where H is the solution of $\mathbb{E}_{\infty} \tau_R(H) = C$ for some pre-determined C . The SR statistic (1.2.6) satisfies the following recurrence:

$$R_n = (1 + R_{n-1}) \cdot \frac{g(y_n)}{f(y_n)}, \quad n \geq 1, R_0 = 0.$$

1.2.2 Evaluating ARL for CUSUM and SR tests

Explicit expressions for $\mathbb{E}_{\infty} \tau_V(H)$ and $\mathbb{E}_{\infty} \tau_R(H)$ are not known. However, they can be numerically obtained by numerically solving particular Fredholm integral equations as proved in [77]. Here it was shown that $\mathbb{E}_{\infty} \tau_V(H)$ and $\mathbb{E}_{\infty} \tau_R(H)$ can be computed by a unified approach for general Markov statistics. Set $H > 0$. For a sufficiently smooth positive valued function ξ and $s \in [0, H]$, let

$$S_n = \xi(S_{n-1}) \cdot \frac{g(y_n)}{f(y_n)}, \quad n \geq 1, S_0 = s \in [0, H]$$

be a Markov detection statistic with stopping rule

$$\tau_S(H) := \inf\{n \geq 1 : S_n > H\}.$$

Let $\phi(s) = \mathbb{E}_{\infty}(\tau_S(H))$ be the ARL (note the dependence on $S_0 = s$) and set $F(x) = \Pr_{\infty}(g(y_1)/f(y_1) \leq x)$. Then $\phi(s)$ is the solution of the following Fredholm integral equation:

$$\phi(s) = 1 + \int_0^H \phi(x) \left[\frac{d}{dx} F \left(\frac{x}{\xi(s)} \right) \right] dx. \quad (1.2.7)$$

For the CUSUM and SR procedures we have $\xi(s) = \max(1, s)$ and $\xi(s) = 1 + s$, respectively. To solve this integral equation, see [77].

Approximations for ARL of the CUSUM and SR procedures have been specifically developed for the problem of detecting the change in mean of normal random variables. Here we operate under (1.2.5). To approximate ARL for both the CUSUM and SR procedures or to narrow the domain of search and more efficiently numerically solve the Fredholm equation (1.2.7), one could use the following simple approximations developed in [129] and [93] respectively:

$$\mathbb{E}_\infty \tau_V(H) \simeq 2H/(A\kappa^2(A)), \quad (1.2.8)$$

$$\mathbb{E}_\infty \tau_R(H) \simeq H/\kappa(A), \quad (1.2.9)$$

where

$$\kappa(A) = \frac{2}{A^2} \exp \left\{ -2 \sum_{\nu=1}^{\infty} \frac{1}{\nu} \Phi \left(-\frac{A}{2} \sqrt{\nu} \right) \right\} \text{ and } \Phi(x) = \int_{-\infty}^x f(y) dy.$$

The approximations in (1.2.8) and (1.2.9) are extremely accurate. In Table 1.1, one can observe the high accuracy of approximation (1.2.8) for different thresholds H . In fact, (1.2.9) is remarkably accurate and frequently leads to exact values of ARL. The only slight inconvenience of both approximations is the numerical evaluation required to compute $\kappa(A)$. This quantity is frequently approximated, see [117, Ch. IV], with $\kappa(A) \simeq \exp(-\rho \cdot A)$, where the constant ρ is defined later in (1.3.16) but can be approximated to three decimal places by $\rho \simeq 0.583$. Using this approximation for κ in (1.2.8) and (1.2.9) still results in excellent approximations. In this table, $\mathbb{E}_\infty \tau_V(H)$ has been approximated Monte Carlo simulations with 100,000 repetitions.

H	9.32	17.33	80.65	159.35	788.00
$\mathbb{E}_\infty \tau_V(H)$	50	100	500	1000	5000
Approximation (1.2.8)	59	110	513	1014	5018
(1.2.8) with $\kappa(A) \simeq \exp(-\rho \cdot A)$	60	111	517	1023	5058

Table 1.1: Approximations for $\mathbb{E}_\infty \tau_V(H)$ with $A = 1$.

1.2.3 Optimality criteria

Denote by $\Delta(C)$ the set of all stopping times of change-point procedures with ARL of at least C . More precisely, $\Delta(C) := \{\tau : \mathbb{E}_\infty \tau \geq C\}$, $C > 1$, where $\tau = \tau(H)$ is a stopping time for a sequential change-point procedure. A common criterion for comparing change-point procedures when $l = \infty$ is the supremum Average Delay to Detection (ADD) introduced by Pollak [92]. Define $ADD_\nu(\tau) := \mathbb{E}_\nu(\tau - \nu | \tau > \nu)$. Then

$$SADD(\tau) := \sup_{0 \leq \nu < \infty} ADD_\nu(\tau). \quad (1.2.10)$$

An optimal change-point procedure would satisfy $SADD(\tau_{opt}) = \inf_{\tau \in \Delta(C)} SADD(\tau)$ for all $C > 1$. Finding an optimal procedure for this criterion is very difficult, where

in general only asymptotic optimality as $C \rightarrow \infty$ (low false alarm rate) is known [92]. Another popular criterion is the worst-case minimax scenario of Lorden [68] defined as

$$\mathcal{L}(\tau) := \sup_{\nu \geq 0} \text{ess sup } \mathbb{E}_\nu[(\tau - \nu)^+ | y_1, y_2, \dots, y_\nu]. \quad (1.2.11)$$

In other words, the conditional ADD is first maximized over all possible trajectories of observations up to the change-point and then over the change-point. We refer the reader to Section 6.3.3 of [131] for further discussions regarding this criterion. Asymptotic optimality (as $C \rightarrow \infty$) of the CUSUM chart of Page was proved in [68]. It was subsequently proved in [75] that the CUSUM chart of Page is in fact optimal under this criterion for every $C > 1$. We refer the reader to Section 6.3.3 of [131] for further discussions regarding this criterion.

The SR procedure is optimal for every $C > 1$ under the Stationary Average Delay to Detection (STADD) criterion. The STADD criterion rewards detection procedures that detect the change as quickly as possible, at the expense of raising many false alarms (using a repeated application of the same stopping rule). Formally, the STADD criterion is defined as follows. Let $\tau_1, \tau_2 \dots$ be a sequence of independent copies of the stopping time τ . Let $T_j = \tau_1 + \tau_2 + \dots + \tau_j$ be the time the j^{th} alarm is raised. Let $I_\nu = \min\{j > 1 : T_j > \nu\}$; this is the index of the first alarm which is not false after $I_\nu - 1$ false alarms. Then

$$STADD(\tau) := \lim_{\nu \rightarrow \infty} \mathbb{E}_\nu[T_{I_\nu} - \nu].$$

It was proved in [96] that the STADD criterion is equivalent to the Relative Integral Average Detection Delay (RIADD) measure (see [77]):

$$RIADD(\tau) := \frac{\sum_{\nu=0}^{\infty} \mathbb{E}_\nu[(\tau - \nu)^+]}{\mathbb{E}_\infty[\tau]}.$$

It is discussed in [77] for both CUSUM and the Shiryaev–Roberts procedure, Lorden’s essential supremum measure (1.2.11) and Pollak’s supremum measure $SADD$ defined in (1.2.10) are attained at $\nu = 0$, that is:

$$\mathcal{L}(\tau_V(H)) = SADD(\tau_V(H)) = \mathbb{E}_0\tau_V(H), \quad \mathcal{L}(\tau_R(H)) = SADD(\tau_R(H)) = \mathbb{E}_0\tau_R(H).$$

Similarly to the computation of $\mathbb{E}_\infty\tau_V(H)$ and $\mathbb{E}_\infty\tau_R(H)$, to obtain $\mathbb{E}_0\tau_V(H)$ and $\mathbb{E}_0\tau_R(H)$ one can numerically solve a Fredholm equation. Instead of setting $\phi(s) = \mathbb{E}_\infty(\tau(H))$, let $\phi(s) = \mathbb{E}_0(\tau(H))$. Also set $F(x) = \Pr_0(g(y_1)/f(y_1) \leq x)$. Then from [77], $\phi(s)$ is the solution of the Fredholm integral equation given in (1.2.7). The computation of STADD requires solving a slightly more difficult integral equation, see [77] for more discussions.

For the Gaussian example considered in (1.2.5), the findings of [77] indicate that for small values of A , say $A = 0.01$, the CUSUM noticeably outperforms the SR procedure under Lorden’s criterion. Vice versa, the SR procedure noticeably outperforms CUSUM under the STADD framework. When the change in A becomes large, say $A = 1$, the benefits a procedure has over the other diminishes.

1.3 Transient change in distributions

In this section, we assume $1 \leq l < \infty$ and therefore study procedures aimed at detecting a transient change in distributions. Suppose y_1, y_2, \dots, y_n have been sampled. The log likelihood ratio for testing \mathbb{H}_∞ against \mathbb{H}_ν is

$$\Gamma_{\nu, \nu+l} := \log \Lambda_{\nu, \nu+l} = \sum_{i=\nu+1}^{\nu+l} \log \frac{g(y_i)}{f(y_i)}. \quad (1.3.1)$$

1.3.1 A collection of procedures

For l unknown, the log likelihood ratio statistic is obtained by maximising (1.3.1) over all possible change point locations ν and transient change lengths:

$$K_n := \max_{0 \leq \nu < \nu+l \leq n} \Gamma_{\nu, \nu+l}, \quad (1.3.2)$$

with the stopping rule

$$\tau_K(H) := \inf\{n \geq 1 : K_n > H\}.$$

Note that in (1.3.2), we are maximising over l too. If there are no nuisance parameters present in f and g that require estimation, the statistic (1.3.2) satisfies the recursive property:

$$K_n = \max\{K_{n-1}, \max_{0 \leq \nu \leq n-1} \Gamma_{\nu, n}\}, \quad K_0 = 0. \quad (1.3.3)$$

For large n , the statistic (1.3.2) is very expensive to compute despite the recursive property given in (1.3.3). This is because in $\max_{0 \leq \nu \leq n-1} \Gamma_{\nu, n}$, one has to maximise over all possible change-point locations which is expensive for large n . For offline change-point problems, this large computational expense may be an inconvenience but it is not a fundamental problem as time is often not an issue. However, for online procedures that require calculations in real time, the statistic K_n is not practical. The assumption of no prior knowledge about the transient change length is unlikely. One can imagine that some knowledge about the length of transient change is likely, for example it may be bounded $l_0 \leq l \leq l_1$. From here on, this assumption will be made. The log likelihood ratio statistic is:

$$Z_n = Z_n(l_0, l_1) := \max_{\substack{0 \leq \nu < \nu+l \leq n \\ l_0 \leq l \leq l_1}} \Gamma_{\nu, \nu+l}, \quad (1.3.4)$$

with the stopping rule

$$\tau_Z(H) := \inf\{n \geq l_1 : Z_n(l_0, l_1) > H\}. \quad (1.3.5)$$

If no nuisance parameters require estimation, the statistic Z_n satisfies the following recursive property:

$$Z_n = \max\left\{Z_{n-1}, \max_{n-l_1 \leq \nu \leq n-l_0} \Gamma_{\nu, n}\right\}, \quad Z_{l_1} = \max_{\substack{0 \leq \nu < \nu+l \leq l_1 \\ l_0 \leq l \leq l_1}} \Gamma_{\nu, \nu+l}. \quad (1.3.6)$$

This is much easier to compute than (1.3.2) for n large. If we make the additional assumption that l is known exactly and is completely contained within the sample of size n , i.e. $\nu + l \leq n$, then the MOSUM statistic is obtained by maximising (1.3.1) over all valid change-point locations ν :

$$M_n := \max_{0 \leq \nu \leq n-l} \Gamma_{\nu, \nu+l}, \quad (1.3.7)$$

The MOSUM statistic can be obtained by setting $l_0 = l_1 = l$ in (1.3.4). For this reason, the statistic Z_n can be called the generalised MOSUM procedure.

The stopping rule associated with MOSUM procedure is:

$$\tau_M(H) := \inf\{n \geq l : M_n > H\}.$$

In what follows, we define the MOSUM test for a general window length L , with L a fixed positive integer. The reason for doing so is as follows. The MOSUM test will be the main change-point algorithm theoretically studied in this chapter and therefore the main algorithm studied in Part one of this thesis. We will be interested in studying quantities like the loss of power, when incorrect information is provided for the true l . Hence L could now be different to l . Results for the likelihood ratio test can still be obtained by setting $L = l$. Define the moving sums

$$S_{n,L} := S_{n,L,L} = \sum_{j=n+1}^{n+L} \log \frac{g(y_i)}{f(y_i)} \quad (n = 0, 1, \dots).$$

Then the stopping rule $\tau_M(H)$ for a given window length L can be expressed as

$$\tau_M(H) = \tau_{S,L}(H) + L, \quad \text{where } \tau_{S,L}(H) := \inf\{n \geq 0 : S_{n,L} > H\} \quad (1.3.8)$$

and therefore $\mathbb{E}_\infty \tau_M(H) = \mathbb{E}_\infty \tau_{S,L}(H) + L$. The moving sum $S_{n,L}$ provided the motivation for the MOSUM name.

For the transient change-point problem, the false alarm risk can be measured through ARL. However, this is not the only approach taken in the change-point literature. In [39] and [61], the false alarm risk is measured through:

$$\sup_{k \geq 1} \Pr_\infty(k \leq \tau < k + m_\alpha) \leq \alpha,$$

where τ is a stopping rule, α is your false alarm tolerance (type 1 error) and $\liminf m_\alpha / |\log(\alpha)| > I_g^{-1}$ but $\log m_\alpha = o(|\log \alpha|)$ as $\alpha \rightarrow 0$; recall $I_g = \mathbb{E}_0(\log(g(y_1)/f(y_1)))$. Another alternative to the usual ARL constraint has been proposed in [129, 130]. Here, the suggested criterion is

$$\sup_{k \geq 1} \Pr_\infty(\tau < k + m_\alpha | \tau \geq l) \leq \alpha.$$

From now on, we measure false alarm risk through ARL $\mathbb{E}_\infty \tau$ and simply refer to [39, 61, 129, 130] for more discussions on other approaches. The majority of research has focused on detecting transient changes in the mean of a sequence of Gaussian random variables. The next section is devoted solely to this problem.

1.3.2 Detecting a transient change in Gaussian random variables

Consider the problem of detecting the change in mean of normal random variables. Suppose pre-change observations are i.i.d. $N(\mu, 1)$ random variables and post-change observations are i.i.d. $N(\mu + A, 1)$ for some $A > 0$. The values of μ, l and A may be known or unknown, with μ and A playing the roles of nuisance parameters if unknown. We have

$$f(y) = \frac{1}{\sqrt{2\pi}} \exp(-(y - \mu)^2/2), \quad g(y) = \frac{1}{\sqrt{2\pi}} \exp(-(y - \mu - A)^2/2). \quad (1.3.9)$$

The offline version of this change-point problem is devoted to testing for change-points in a sample of fixed length and has seen significant attention in the past, see [43, 64, 118, 119, 139]. An excellent survey of several statistics aimed at addressing the offline problem can be found in [141]. Despite the fact Z_n defined in (1.3.4) is a generalisation of M_n given in (1.3.7), we will initially discuss recent results for M_n . These results will provide inspiration for addressing the much more complicated problems associated with Z_n .

1.3.2.1 The MOSUM statistic

The MOSUM stopping rule given in (1.3.8) specialised for this Gaussian example is

$$\begin{aligned} \tau'_M(H') &= \tau'_{S,L}(H') + L, \quad \tau'_{S,L}(H') = \inf\{n \geq 0 : S'_{n,L} > H'\} \\ \text{with } S'_{n,L} &= A \sum_{j=n+1}^{n+L} (y_j - \mu - A/2). \end{aligned}$$

Knowledge of A is required to set the ARL constraint $\mathbb{E}_\infty \tau'_M(H') = C$. However, if we consider the stopping rule

$$\begin{aligned} \tau_M(H) &= \tau_{S,L}(H) + L, \quad \tau_{S,L}(H) = \inf\{n \geq 0 : S_{n,L} > H\} \\ \text{with } S_{n,L} &= \sum_{j=n+1}^{n+L} y_j, \end{aligned} \quad (1.3.10)$$

then one can show that $\tau'_M(H') \stackrel{d}{=} \tau_M(H'/A + \mu L + AL/2)$. As a result, the stopping rules $\tau'_M(H')$ and $\tau_M(H)$ are equal in distribution provided with $\mathbb{E}_\infty \tau'_M(H') = \mathbb{E}_\infty \tau_M(H) = C$. The stopping rule $\tau_M(H)$ has the benefit of not requiring knowledge of A to set the ARL constraint. We will refer to the stopping rule $\tau_M(H)$ as the MOSUM test in this Gaussian setting.

The problem of approximating $\mathbb{E}_\infty \tau_{S,L}(H)$ assuming μ is known is a main considered in Chapter 2 (Section 1.3). Here will recall the main steps in the construction and refer to Chapter 2 for the specific details. Define

$$h = \frac{H - \mu L}{\sqrt{L}} \quad \text{so that} \quad H = \mu L + h\sqrt{L}$$

and consider the standardised versions of $S_{n,L}$:

$$\xi_{n,L} := \frac{S_{n,L} - \mathbb{E}_\infty S_{n,L}}{\sqrt{\text{Var}_\infty(S_{n,L})}} = \frac{S_{n,L} - \mu L}{\sqrt{L}}, \quad n = 0, 1, \dots$$

Then the stopping time $\tau_{S,L}(H)$ is equivalent to the stopping rule

$$\tau_\xi(h) := \inf\{n \geq 0 : \xi_{n,L} \geq h\}$$

and hence $\mathbb{E}_\infty \tau_\xi(h) = \mathbb{E}_\infty \tau_{S,L}(H)$.

For any integer $M \geq 0$, the discrete time process $\xi_{0,L}, \xi_{1,L}, \dots, \xi_{M,L}$ is approximated by a continuous time analogue $S(t)$ on $[0, T = M/L]$. The process $S(t)$ is a zero mean and variance one, stationary Gaussian process with correlation function $R(t) = \max\{0, 1 - |t|\}$. The ARL $\mathbb{E}_\infty \tau_\xi(h)$ then has the continuous-time approximation

$$\mathbb{E}_\infty \tau_\xi(h) \cong -L \int_0^\infty s dF_s(h), \quad (1.3.11)$$

where $F_T(h) := \Pr_\infty(S(t) < h \text{ for all } t \in [0, T])$.

Explicit formulas for the probability $F_T(h)$ with $T \leq 1$ were first derived in [122]. Here it was shown for $Z = T/(2 - T)$:

$$\begin{aligned} F_T(h) &= \int_{-\infty}^h \Phi\left(\frac{h(Z+1) - x(-Z+1)}{2\sqrt{Z}}\right) \varphi(x) dx \\ &\quad - \frac{2\sqrt{Z}}{Z+1} \varphi(h) \left[h\sqrt{Z} \Phi(h\sqrt{Z}) + \frac{1}{\sqrt{2\pi}} (\sqrt{2\pi} \varphi(h))^Z \right]. \end{aligned}$$

For $T = 1$ this reduces to

$$F_1(h) = \Phi^2(h) - \varphi(h) [h\Phi(h) + \varphi(h)]. \quad (1.3.12)$$

For $T > 1$, formulae for $F_T(h)$ were first derived in [113]; these expressions take different forms depending on whether or not T is integer. The result of [113, p.949] states that if $T = n$ is a positive integer then

$$F_n(h) = \int_{-\infty}^h \int_{D_x} \det[\varphi(y_i - y_{j+1} + h)]_{i,j=0}^n dy_2 \dots dy_{n+1} dx \quad (1.3.13)$$

where $y_0 = 0, y_1 = h - x, D_x = \{y_2, \dots, y_{n+1} \mid h - x < y_2 < y_3 < \dots < y_{n+1}\}$. For non-integer $T \geq 1$, the exact formula for $F_T(h)$ is even more complex (the integral has the dimension $\lceil 2T \rceil + 1$), see [113, p.950]. For $T = 2$, (1.3.13) yields

$$\begin{aligned} F_2(h) &= \Phi^3(h) - 2h\varphi(h)\Phi^2(h) + \frac{h^2 - 3 + \sqrt{\pi}h}{2} \varphi^2(h)\Phi(h) + \frac{h + \sqrt{\pi}}{2} \varphi^3(h) \\ &\quad + \int_0^\infty \Phi(h - y) \left[\varphi(h + y)\Phi(h - y) - \sqrt{\pi} \varphi^2(h)\Phi(\sqrt{2}y) \right] dy. \end{aligned} \quad (1.3.14)$$

The complicated nature of these expressions for $F_T(h)$ made them impractical for the use in the ARL approximation (1.3.11). One simple yet still very accurate approximation has the form:

$$F_T(h) \simeq F_2(h) [\theta(h)]^{T-2}, \quad \text{where } \theta(h) = F_2(h)/F_1(h) \quad (1.3.15)$$

and the probabilities $F_1(h)$ and $F_2(h)$ are given in (1.3.12) and (1.3.14) respectively. This approximation is derived in Chapter 3 of this thesis (see Approximation 4 in that chapter). Here, $\varphi(x)$ and $\Phi(x)$ are the standard normal density and distribution functions respectively. The intuition behind the form of approximation (1.3.15) is as follows, but more detail can be obtained in Chapter 3. Using laws of conditional probability, we can express $F_T(h)$ as:

$$\begin{aligned} F_T(h) &= \Pr_\infty(S(t) < h \text{ for all } t \in (T-1, T] \mid S(t) < h \text{ for all } t \in [0, T-1]) \\ &\times \Pr_\infty(S(t) < h \text{ for all } t \in (T-2, T-1] \mid S(t) < h \text{ for all } t \in [0, T-2]) \\ &\times \cdots \times \Pr_\infty(S(t) < h \text{ for all } t \in (2, 3] \mid S(t) < h \text{ for all } t \in [0, 2]) \\ &\times F_2(h). \end{aligned}$$

The process $S(t)$ has a short memory of length one, i.e. $S(t)$ and $S(t-1)$ are independent. As a result, under the condition $S(t)$ remains under the barrier h , the process reaches stationary behaviour quickly. Consequently, we make the approximation:

$$\begin{aligned} &\Pr_\infty(S(t) < h \text{ for all } t \in (T-1, T] \mid S(t) < h \text{ for all } t \in [0, T-1]) \\ &\simeq \Pr_\infty(S(t) < h \text{ for all } t \in (T-2, T-1] \mid S(t) < h \text{ for all } t \in [0, T-2]) \\ &\simeq \Pr_\infty(S(t) < h \text{ for all } t \in (1, 2] \mid S(t) < h \text{ for all } t \in [0, 2]) \\ &= \theta(h). \end{aligned}$$

The initial term $F_2(h)$ in (1.3.15) should not be reduced further and exists to allow the process $S(t)$ to reach stationarity.

The approximation given in (1.3.15) applied to (1.3.11) results in the following continuous-time ARL approximation:

$$\mathbb{E}_\infty \tau_\xi(h) \simeq -\frac{L \cdot F_2(h)}{\theta(h)^2 \log(\theta(h))}.$$

This approximation was then corrected in Chapter 2, see Section 2.8, for discrete time to improve results for small L . This amounted to correcting the probabilities $F_1(h)$ and $F_2(h)$ for discrete time; this was performed by specialising results of D. Siegmund; primarily on expected overshoot a discrete time normal random walk has over a threshold. From [117, p. 225], this expected overshoot was computed as

$$\rho := -\pi^{-1} \int_0^\infty \lambda^{-2} \log\{2(1 - \exp(-\lambda^2/2))/\lambda^2\} d\lambda \simeq 0.582597. \quad (1.3.16)$$

Define the probability

$$F_L(h, M) := \Pr_\infty \left(\max_{n=0,1,\dots,M} \xi_{n,L} < h \right).$$

From Section 2.8 Chapter 2 (also [83, p. 18]):

$$\begin{aligned} \mathbb{E}_\infty \tau_{S,L}(H) &= \mathbb{E}_\infty \tau_\xi(h) \simeq -\frac{L \cdot F_L(h, 2L)}{\theta_L(h)^2 \log(\theta_L(h))} \\ &\text{with } \theta_L(h) = \frac{F_L(h, 2L)}{F_L(h, L)}, \end{aligned} \quad (1.3.17)$$

where the probabilities $F_L(h, L)$ and $F_L(h, 2L)$ can be approximated using the formulas (2.4.19) and (2.4.20) of Chapter 2 respectively. Although $\Phi(x)$ is formally a one-dimensional integral, in this chapter and thesis we adopt the convention that it is explicit. This is because $\Phi(x)$ can be easily evaluated by all statistical software. By using this convention, only a one-dimensional integral has to be numerically evaluated for approximating $F_L(h, 2L)$. Tables 1.2 and Tables 1.3 demonstrate that (1.3.17) using the formulas of Chapter 2 is extremely accurate.

h	2	2.25	2.5	2.75	3	3.25	3.5
(1.3.17)	126	217	395	759	1551	3375	7837
$\mathbb{E}_\infty \tau_\xi(h)$	127	218	396	757	1550	3344	7721

Table 1.2: Approximations for $\mathbb{E}_\infty \tau_\xi(h)$ with $L = 10$.

h	2	2.25	2.5	2.75	3	3.25	3.5
(1.3.17)	471	791	1392	2587	5099	10695	23918
$\mathbb{E}_\infty \tau_\xi(h)$	472	792	1397	2588	5085	10749	24131

Table 1.3: Approximations for $\mathbb{E}_\infty \tau_\xi(h)$ with $L = 50$.

For approximating the boundary-crossing probability $F_L(h, M)$ for all M , the discrete time corrected form of (1.3.15) suggests using the approximation

$$F_L(h, M) \simeq F_L(h, 2L) [\theta_L(h)]^{M/L-2}. \quad (1.3.18)$$

This approximation is studied in Chapter 2 (see (2.4.21)), where one approximates $F_L(h, 2L)$ and $\theta_L(h)$ using (2.4.19) and (2.4.20) respectively. For a comprehensive assessment of the accuracy of this approximation, see Chapter 2.

1.3.2.2 The stopping rule $\tau_Z(H)$

Here, we assume l is not known exactly but can be bounded between l_0 and l_1 . We will initially assume μ and A are known. The stopping rule given in (1.3.5) specialised for this Gaussian example is tantamount to:

$$\tau_Z(H) = \inf \left\{ n \geq l_1 : \max_{\substack{0 \leq \nu < \nu+l \leq n \\ l_0 \leq l \leq l_1}} Z_n(l_0, l_1) \right\}, \quad (1.3.19)$$

$$\text{with } Z_n = Z_n(l_0, l_1) = A \sum_{j=\nu+1}^{\nu+l} \left(y_j - \mu - \frac{A}{2} \right)$$

Using the recursive property outlined in (1.3.6), for $n > l_1$ the statistic Z_n satisfies:

$$Z_n = \max \{ Z_{n-1}, S_{n, l_0, l_1} \}, \quad \text{with } S_{n, l_0, l_1} := \max_{n-l_1 \leq \nu \leq n-l_0} A \sum_{j=\nu+1}^n \left(y_j - \mu - \frac{A}{2} \right)$$

$$\text{and } Z_{l_1} = \max_{\substack{0 \leq \nu < \nu+l \leq l_1 \\ l_0 \leq l \leq l_1}} A \sum_{j=\nu+1}^{\nu+l} \left(y_j - \mu - \frac{A}{2} \right).$$

The short memory of the MOSUM statistic is paramount to the form of the approximation given in (1.3.18). This short memory is also present within the generalised moving sum statistic Z_n if one considers its recursive definition above. After the initialising value Z_{l_1} , Z_n essentially becomes a moving sum process given by S_{n,l_0,l_1} . The process $\{S_{n,l_0,l_1}\}$ exhibits a short memory, where dependence between two values is lost after l_1 observations i.e. S_{n,l_0,l_1} and S_{n+l_1,l_0,l_1} are independent for all n . This suggests that stationary behaviour of the combined process $(Z_{l_1}, \{S_{n,l_0,l_1}\})$, under the condition of not crossing the barrier H , should be attained quickly. One would then anticipate that the form of approximations (1.3.17) and (1.3.18) would also be suitable when applied to Z_n . For $M \geq 0$, introduce the probability:

$$F_{l_0,l_1}(H, M) := \Pr_{\infty}(Z_{M+l_1} < H) = \Pr_{\infty}\{Z_{l_1} < H, S_{j,l_0,l_1} < H \forall j = l_1 + 1, \dots, l_1 + M\}.$$

Then one would expect (as an educated guess) the following approximation to be accurate:

$$F_{l_0,l_1}(H, M) \simeq F_{l_0,l_1}(H, 2l_1) [\theta_{l_1}(H)]^{M/l_1-2} \text{ with } \theta_{l_1}(H) = \frac{F_{l_0,l_1}(H, 2l_1)}{F_{l_0,l_1}(H, l_1)}, \quad (1.3.20)$$

$$\mathbb{E}_{\infty}\tau_Z(H) \simeq l_1 - \frac{l_1 \cdot F_{l_0,l_1}(H, 2l_1)}{[\theta_{l_1}(H)]^2 \log(\theta_{l_1}(H))}. \quad (1.3.21)$$

Unfortunately, the probability $F_{l_0,l_1}(H; M)$ is complex and to the author's knowledge no formula or approximations are known. The probabilities $F_{l_0,l_1}(H; 2l_1)$ and $F_{l_0,l_1}(H; l_1)$ can be approximated via simulations; this is not too cumbersome as at most $3l_1$ random variables need to be simulated at each iteration. As commonly $\mathbb{E}_{\infty}\tau_Z(H) = C$ with C large, the right tail of the distribution of the random variable $\max\{Z_{l_1}, \max_{j=0 \dots M} S_{j,l_0,l_1}\}$ is of the most interest. Large deviation theory, see [118, 139, 140], could be used to approximate the right tail of this distribution, however numerical results indicate approximations of these kind would not be accurate enough for general l_0 and l_1 (those that are not astronomically large). If the prior knowledge that $1 \leq l \leq l_1$ is known, and an explicit formula to approximate $F_{1,l_1}(H; M)$ or $\mathbb{E}_{\infty}\tau_Z(H)$ is desired, the following heuristic argument could be used. Using inspiration from [43], a continuous time analogue of the probability $F_{1,l_1}(H; M)$ that allows for the application of existing large deviation results is:

$$\Pr \left\{ \max_{\substack{0 \leq s < t \leq M+l_1 \\ 0 \leq t-s \leq l_1}} \left[W(t) - W(s) - \frac{A}{2}(t-s) \right] < \frac{H}{A} \right\}, \quad (1.3.22)$$

where $W(t)$, $0 \leq t < \infty$, is standard Brownian motion. Ideally, a large deviation approximation for (1.3.22) should be computed explicitly. However, for $M = l_1$, $M = 2l_1$ and A large, say $A \geq 1$, simulation studies indicate that the additional maximisation constraint in (1.3.22) of $0 < t - s < l_1$ has very little influence on this probability. If this constraint is ignored, the following large deviation result of [43] can be applied.

Lemma 1.3.1 *Suppose $\gamma > 0$, $m \rightarrow \infty$ and $u \rightarrow \infty$ such that $m\gamma u^{-1}$ is some fixed number in $(1, \infty)$. Then*

$$\Pr \left\{ \max_{0 \leq s < t \leq m} [W(t) - W(s) - \gamma(t - s)] > u \right\} = [2\gamma(m\gamma - u) + 3 + o(1)] \exp(-2\gamma u).$$

To subsequently correct this result for discrete time, it is recommended in [43] to increase the barrier H by 2ρ , where ρ is defined in (1.3.16). This results in the approximations

$$\begin{aligned} F_{1,l_1}(H, l_1) &\simeq 1 - (A(Al_1 - H/A - 2\rho) + 3) \exp\{-A(H/A + 2\rho)\} \\ F_{1,l_1}(H, 2l_1) &\simeq 1 - (A(3Al_1/2 - H/A - 2\rho) + 3) \exp\{-A(H/A + 2\rho)\}. \end{aligned}$$

As a result, using the approximations given in (1.3.20) and (1.3.21):

$$F_{1,l_1}(H, M) \simeq 1 - (A(3Al_1/2 - H/A - 2\rho) + 3) \exp\{-A(H/A + 2\rho)\} \left[\hat{\theta}_{l_1}(H) \right]^{M/l_1 - 2} \quad (1.3.23)$$

with

$$\hat{\theta}_{l_1}(H) = \frac{1 - (A(3Al_1/2 - H/A - 2\rho) + 3) \exp\{-A(H/A + 2\rho)\}}{1 - (A(Al_1 - H/A - 2\rho) + 3) \exp\{-A(H/A + 2\rho)\}}.$$

Also

$$\mathbb{E}_\infty \tau_Z(H) \simeq l_1 - \frac{l_1 [1 - (A(3Al_1/2 - H/A - 2\rho) + 3) \exp\{-A(H/A + 2\rho)\}]}{[\hat{\theta}_{l_1}(H)]^2 \log(\hat{\theta}_{l_1}(H))}. \quad (1.3.24)$$

The accuracy of the approximation in (1.3.20) is demonstrated in Figure 1.1 for different l_0, l_1, M and A as a function of H . In this approximation, $F_{l_0, l_1}(H, 2l_1)$ and $F_{l_0, l_1}(H, l_1)$ have been approximated using Monte Carlo simulations with 50,000 repetitions. In this figure, the probability $F_{l_0, l_1}(H; M)$ is depicted with a thick dashed black line and is obtained from 50,000 simulations. The approximation in (1.3.20) is depicted with a solid blue line. From this figure, the high accuracy of approximation (1.3.20) is clearly demonstrated. In Figures 1.2-1.3, we assess the accuracy of the approximation in (1.3.23). In these figures, for $A = 1$ and various M , the probability $F_{1, l_1}(H; M)$ is depicted with a thick dashed black line whereas the approximation provided in (1.3.23) is shown with a solid red line. The number present on the figure is used to show the value of l_1 used. From these figures, we see for large H the approximation in (1.3.23) is adequate. In Tables 1.4-1.5 the accuracy of the approximations provided in (1.3.21) and (1.3.24) are assessed for different H . We see the approximation in (1.3.21) is extremely accurate for all H . For large H , the approximation in (1.3.24) is fairly accurate and has the benefit of explicit evaluation. For small H and small A , the accuracy of (1.3.24) should deteriorate. In these tables, $\mathbb{E}_\infty \tau_Z(H)$ has been approximated using Monte Carlo simulations with 10,000 repetitions and $F_{l_0, l_1}(H, l_1)$ and $F_{l_0, l_1}(H, 2l_1)$ used in approximation (1.3.21) have been approximated using 50,000 repetitions.

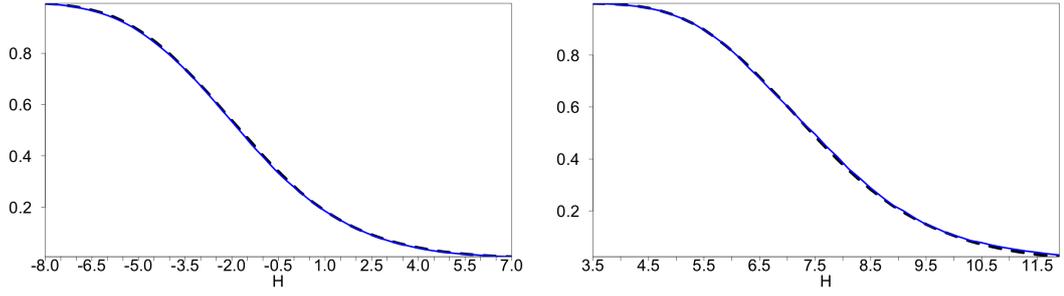


Figure 1.1: Empirical probabilities of reaching the barrier H (dashed black) and approximation (1.3.20) (solid blue). Left: $A = 1$, $M/l_1 = 4$ with $l_0 = 25$ and $l_1 = 50$. Right: $A = 0.5$, $M/l_1 = 25$ with $l_0 = 10$ and $l_1 = 20$.

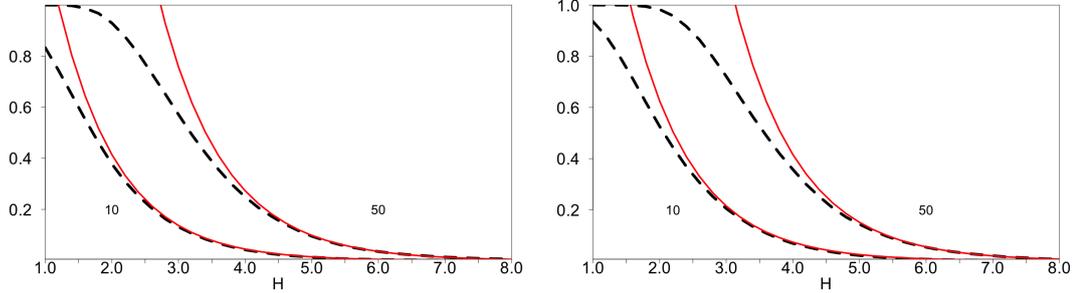


Figure 1.2: Empirical probabilities of reaching the barrier H (dashed black) and corresponding versions of approximation (1.3.23) (solid red). Left: $A = 1$, $M/l_1 = 1$ with (a) $l_1 = 10$ and (b) $l_1 = 50$. Right: $A = 1$, $M/l_1 = 2$ with (a) $l_1 = 10$ and (b) $l_1 = 50$.

H	-5	-4.5	-4	-3.5	-3	-2.5	-2
(1.3.21)	126	145	166	196	228	268	319
$\mathbb{E}_\infty \tau_Z(H)$	127	144	167	194	229	272	323

Table 1.4: Approximations for $\mathbb{E}_\infty \tau_Z(H)$ with $l_0 = 25$, $l_1 = 50$, $A = 1$.

H	2	2.25	2.5	2.75	3	3.25	3.5
(1.3.24)	30	42	59	81	111	148	195
(1.3.21)	41	53	70	91	120	156	205
$\mathbb{E}_\infty \tau_Z(H)$	41	54	70	91	120	157	207

Table 1.5: Approximations for $\mathbb{E}_\infty \tau_Z(H)$ with $l_0 = 1$, $l_1 = 10$, $A = 1$.

1.3.2.3 The presence of nuisance parameters

Here we briefly consider statistics aimed at detecting a transient change when certain nuisance parameters require estimation. The brevity of this discussion is because (in the authors opinion) in practice for online change-point problems, the behaviour of the time series under the null hypothesis of no change-point is often observed for a

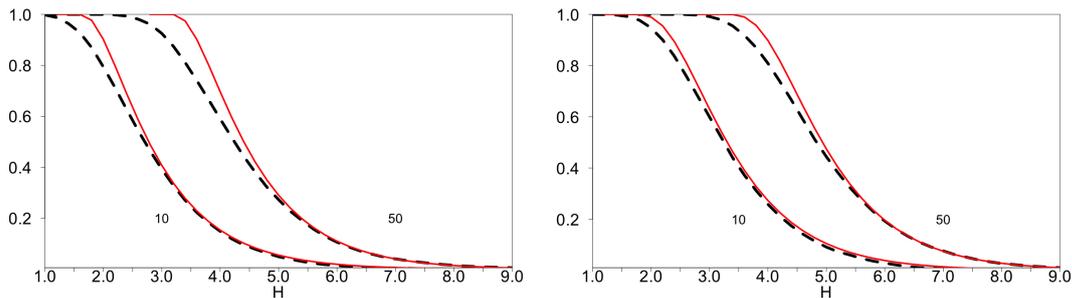


Figure 1.3: Empirical probabilities of reaching the barrier H (dashed black) and corresponding versions of approximation (1.3.23) (solid red). Left: $A = 1$, $M/l_1 = 5$ with (a) $l_1 = 10$ and (b) $l_1 = 50$. Right: $A = 1$, $M/l_1 = 10$ with (a) $l_1 = 10$ and (b) $l_1 = 50$.

lengthy period of time. This allows for the accurate estimation of certain nuisance parameters and they can therefore be assumed known. The counterargument to this is that small trend changes or small undetectable changes may contaminate estimates. For the situation of small trend changes, the use of singular spectrum analysis (SSA) could be used to extract the trend allowing for the study of only the residuals, see Section 1.3.5 where an approach similar to this is discussed. Many of the following statistics appear in some form in [141] when addressing the offline change-point problem, and a number of approximations for the false alarm error are provided. The log likelihood ratio given in (1.3.1), where f and g are given in (1.3.9), is

$$\Gamma_{\nu, \nu+l} = A \sum_{j=\nu+1}^{\nu+l} \left(y_j - \mu - \frac{A}{2} \right).$$

Using motivation from [64], if μ is unknown, l is unknown but bounded $l_0 \leq l \leq l_1$ and A is known, then one can replace μ with its maximum likelihood estimator under H_∞ ; $\hat{\mu} := \sum_{i=1}^n y_i/n$ to obtain:

$$Z_n^1 := \max_{\substack{0 \leq \nu < \nu+l \leq n \\ l_0 \leq l \leq l_1}} A \sum_{j=\nu+1}^{\nu+l} \left(y_j - \hat{\mu} - \frac{A}{2} \right).$$

In [118], μ was replaced with its average over the null and alternative hypotheses to obtain the true likelihood ratio statistic:

$$Z_n^2 := \max_{\substack{0 \leq \nu < \nu+l \leq n \\ l_0 \leq l \leq l_1}} A \sum_{j=\nu+1}^{\nu+l} \left(y_j - \hat{\mu} - \frac{A}{2} \left(1 - \frac{l}{n} \right) \right).$$

If μ and A are both unknown, the square root of the log likelihood ratio statistic is:

$$Z_n^3 := \max_{\substack{0 \leq \nu < \nu+l \leq n \\ l_0 \leq l \leq l_1}} \frac{\left(\sum_{j=\nu+1}^{\nu+l} y_j \right) - l \hat{\mu}}{\sqrt{l \left(1 - \frac{l}{n} \right)}}.$$

The statistic Z_n^3 is also studied in [106, p. 497] for detecting a transient change in mean of random variables in an offline setting. Here the statistic is formulated without likelihood arguments and is therefore used when y_j are not necessarily Gaussian. In [106], the authors view the change-point statistic as a discretizations of some Holder norms or semi-norms allowing them to obtain limiting distributions under the null hypothesis of no change in mean.

Instead of testing for the existence of a single transient change in the mean, the problem of detecting multiple changes in the means of i.i.d. random variables has been studied in [23]. Here, a MOSUM-like statistic is used for detecting any possible number of changes in a sample of fixed length n and the values of the mean after the change-point do not necessarily have to be known. This can be seen as a significant generalisation of the problems considered in this chapter and in [141], if one considers only the offline setting. The statistic studied in [23] is proportional to the following quantity:

$$\max_{L \leq \nu \leq n-L} |T_{\nu,n}(L)|, \quad (1.3.25)$$

with

$$T_{\nu,n}(L) = \frac{1}{\sqrt{2L}} \left(\sum_{i=\nu+1}^{\nu+L} y_i - \sum_{i=\nu-L+1}^{\nu} y_i \right).$$

The statistic $T_{\nu,n}(L)$ has a simple interpretation of comparing at every time point $L \leq \nu \leq n-L$ the mean of the subsample $y_{\nu-L+1}, \dots, y_{\nu}$ with the mean of the subsample $y_{\nu+1}, \dots, y_{\nu+L}$. Naturally, a large difference between the two means (the sign is irrelevant because of the absolute value in (1.3.25)) would indicate a change at this point. As mentioned in [23], at a point ν this statistic is similar to the likelihood ratio statistic for the sample $y_{\nu+1}, \dots, y_{\nu+L}$ at the potential change-point ν . The asymptotic behaviour of a normalised form of the statistic in (1.3.25) as $n \rightarrow \infty$ is given Theorem 2.1 of [23] (here we are ignoring a number of technical details) and is shown to follow a Gumbel extreme value distribution. The Gumbel extreme value distribution for the MOSUM test given in (1.3.10) will be mentioned in Section 2.7 of Chapter 2.

For the statistics considered in this section, it is not obvious how one can translate the offline change-point results of [23, 43, 64, 106, 118, 119, 139] to address the online change-point problem in the presence of nuisance parameters.

1.3.3 Optimality criteria

For online detection of transient changes, optimality criteria like (1.2.10) and (1.2.11) do not have much meaning as the change in distributions is not permanent (signal can be missed). Instead, optimality involving the maximisation of the probability of detection under a constraint on the false alarm risk is more appropriate, see [3, 41]. One could use a worst-case criterion of the form:

$$\inf_{\nu} \Pr_{\nu} \{ \tau(H) - \nu < T \mid \tau(H) > \nu \}, \quad (1.3.26)$$

where $T > 1$ is the maximum length of time after the change-point occurs that it must be detected; this is problem specific and is therefore chosen by the user. By imposing the condition of a long run with no false alarms, another possible criterion is

$$\lim_{\nu \rightarrow \infty} \Pr_{\nu} \{ \tau(H) - \nu < T \mid \tau(H) > \nu \}. \quad (1.3.27)$$

Using ARL as the measure of false alarm risk, a stopping rule $\tau \in \Delta(C)$ is then optimal for a given C if it maximises (1.3.26) or (1.3.27); recall $\Delta(C) = \{ \tau : \mathbb{E}_{\infty} \tau \geq C \}$, $C > 1$.

1.3.3.1 MOSUM procedure

For the MOSUM procedure given in (1.3.10), the quantity (1.3.27) is the focus of study in Chapter 5 and builds on the continuous time results that will be demonstrated in Chapter 4. For $T = l + L$, the quantity (1.3.27) is equivalent to:

$$\mathcal{P}_S(H, A, L) := \lim_{\nu \rightarrow \infty} \Pr_{\nu} \{ S_{n,L} > H \text{ for some } n \in [\nu' + 1, \nu + l - 1] \mid \tau_{S,L}(H) > \nu' \}. \quad (1.3.28)$$

with $\nu' := \nu - L$.

Formally, we require $\nu \rightarrow \infty$ in (1.3.28). This is to ensure that the sequence of moving sums $\{S_{n,L}\}_n$ reaches the stationary behaviour under the null hypothesis and given that we have not crossed the threshold H . However, as shall be discussed in Chapters 4 and 5, this stationary regime is reached very quickly and in all approximations it is enough to only require $\nu \geq 2L$.

The reasoning behind the choice $T = l + L$ is as follows. Assume \mathbb{H}_{ν} with $\nu < \infty$, and that ν is suitably large. If the barrier H is reached for any sum $S_{n,L}$ with $n \leq \nu'$ then, since there are no parts of the signal in the sums $S_{0,L}, \dots, S_{\nu',L}$, we classify the event of reaching the barrier as a false alarm. Each one of the sums $S_{\nu'+1,L}, \dots, S_{\nu+l-1,L}$ has mean larger than $L\mu$ as it contains at least a part of the signal. Reaching the barrier H by any of these sums will be classified as a correct detection of the signal. If neither of these sums reaches H , then we say that we failed to detect the signal and further events when $S_{n,L} \geq H$ with $n \geq \nu + l$ will again be classified as false alarms. In Figure 1.4 we display the values $\mathbb{E}_{\nu} S_{n,L}$ as a function of n .

Define the function

$$Q(n; A, L, \nu') := \begin{cases} 0 & \text{for } n \leq \nu' \text{ or } n \geq \nu + l \\ A(n - \nu') & \text{for } \nu' < n \leq \nu' + \min(l, L) \\ A \min(l, L) & \text{for } \nu' + \min(l, L) < n \leq \nu' + \max(l, L) \\ A(L + l + \nu' - n) & \text{for } \nu' + \max(l, L) < n \leq \nu + l - 1. \end{cases}$$

Then Figure 1.4 is also a plot of $\mu L + Q(n; A, L, \nu')$. By subtracting $\mathbb{E}_{\nu} S_{n,L}$ from the threshold H and standardising the random variables $S_{n,L}$ the power of the test given in (1.3.28) can be expressed in terms of probability under \mathbb{H}_{∞} :

$$\mathcal{P}_{\xi}(h, A, L) := \lim_{\nu \rightarrow \infty} \Pr_{\infty} \left\{ \xi_{n,L} > h - \frac{Q(n; A, L, \nu')}{\sigma \sqrt{L}} \text{ for some } n \in [\nu' + 1, \nu + l - 1] \mid \tau_{\xi}(h) > \nu' \right\},$$

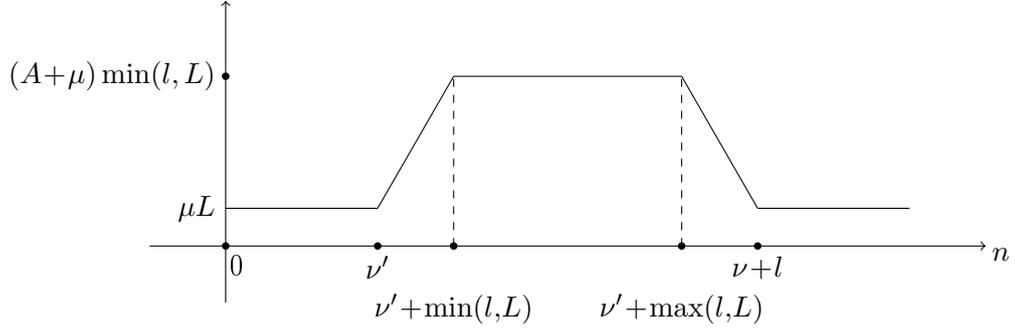


Figure 1.4: $\mathbb{E}_\nu S_{n,L}$ as a function of n

where $\mathcal{P}_S(H, A, L) = \mathcal{P}_\xi(h, A, L)$. Recall the relation $H = \mu L + h\sqrt{L}$. To approximate $\mathcal{P}_\xi(h, A, L)$, the approach taken in Chapter 4 (and [84]) was similar to the approach taken to approximate ARL in Section 1.3.2.1. The approach is as follows. We firstly approximate the problem in the continuous-time setting and compute probabilities for the Gaussian process $S(t)$. Then, use the results of D. Siegmund to correct the continuous time probability for discrete time. Fix $\gamma = A\sqrt{L}$, $\kappa = \nu'/L$, $\lambda = l/L$ and define the function

$$Q(t; \gamma, \kappa, \lambda) = \begin{cases} 0 & \text{for } t \leq \kappa \text{ or } t \geq \kappa + 1 + \lambda. \\ \gamma(t - \kappa) & \text{for } \kappa < t \leq \kappa + \min(1, \lambda) \\ \gamma \min(1, \lambda) & \text{for } \kappa + \min(1, \lambda) < t \leq \kappa + \max(1, \lambda) \\ \gamma(1 + \lambda + \kappa - t) & \text{for } \kappa + \max(1, \lambda) < t \leq \kappa + 1 + \lambda. \end{cases}$$

The diffusion approximation for the power of the test is

$$\mathcal{P}(h, A) := \tag{1.3.29}$$

$$\lim_{\kappa \rightarrow \infty} \Pr_\infty \{S(t) > h - Q(t; \gamma, \kappa, \lambda) \text{ for some } t \in [\kappa, \kappa + 1 + \lambda] \mid \tilde{\tau}(h) > \kappa\},$$

where $\tilde{\tau}(h) = \inf\{t > 0 : S(t) > h\}$. We refer to Lemma 5.4.1 in Chapter 4 for more details about this approach. That is, by assuming $L \rightarrow \infty$, we make the approximation

$$\mathcal{P}_\xi(h, A, L) \cong \mathcal{P}(h, A).$$

The complexity of computation of the diffusion approximation $\mathcal{P}(h, A)$ and its discrete-time corrected version depends on the choice of L in comparison to l . Here, we will only consider the scenario of $\lambda = l/L = 1$ which corresponds to the case of l known at the MOSUM construction stage. The two other cases of $\lambda > 1$ and $\lambda < 1$ are studied in Chapter 4.

For $\lambda = 1$, the diffusion approximation for $\mathcal{P}_\xi(h, A, L)$ given in (1.3.29) reduces to

$$\mathcal{P}(h, A) = \tag{1.3.30}$$

$$\lim_{\kappa \rightarrow \infty} \Pr_\infty \{S(t) \geq h - Q(t; \gamma, \kappa) \text{ for some } t \in [\kappa, \kappa + 2] \mid \tilde{\tau}(h) > \kappa\},$$

where $Q(t; \gamma, \kappa) = \gamma \max\{0, 1 - |t - (\kappa + 1)|\}$. The barrier $h - Q(t; \gamma, \kappa)$ is depicted in Figure 1.5.

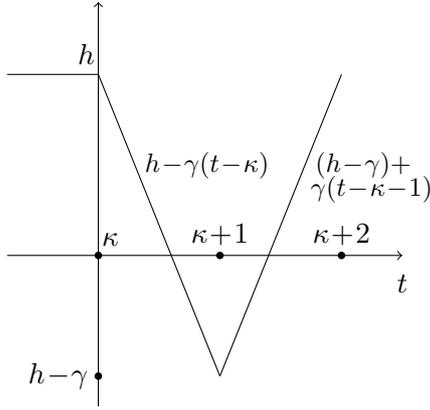


Figure 1.5: Barrier $h - Q(t; \gamma, \kappa)$ for $\lambda = 1$.

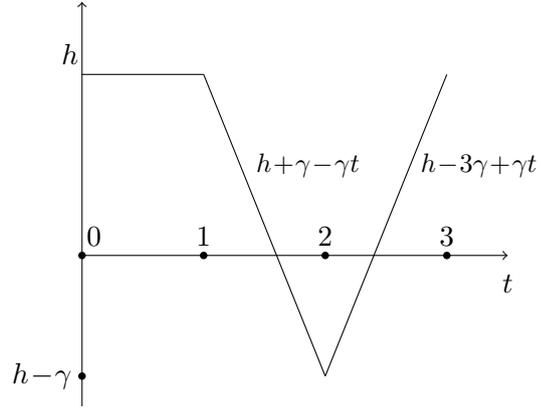


Figure 1.6: Barrier $B(t; h, 0, -\gamma, \gamma)$.

The probability (1.3.30) is considered in Chapter 4, where approximations accurate to more than 4 decimal places are developed. Define the following two conditional probabilities:

$$F_{h,0}(1|x) := \Pr_{\infty}(S(t) < h \text{ for all } t \in [0, 1] \mid S(0) = x),$$

$$F_{h,0,-\gamma,\gamma}(3|x) := \Pr_{\infty}(S(t) < B(t; h, 0, -\gamma, \gamma) \text{ for all } t \in [0, 3] \mid S(0) = x),$$

where the barrier $B(t; h, 0, -\gamma, \gamma)$ is defined as

$$B(t; h, 0, -\gamma, \gamma) = \begin{cases} h, & 0 \leq t \leq 1 \\ h - \gamma(t - 1), & 1 < t \leq 2 \\ h - \gamma + \gamma(t - 2), & 2 < t \leq 3 \\ 0 & \text{otherwise,} \end{cases}$$

and is depicted in Figure 1.6. From (4.5.7) in Chapter 4 we obtain

$$\mathcal{P}(h, A) \cong 1 - \frac{F_{h,0,-\gamma,\gamma}(3|0)}{F_{h,0}(1|0)}, \quad (1.3.31)$$

where

$$F_{h,0}(1|x) = \Phi(h) - \exp(-(h^2 - x^2)/2) \Phi(x)$$

and

$$F_{h,0,-\gamma,\gamma}(3|x) = \frac{e^{\gamma^2/2}}{\varphi(x)} \int_{-x-h}^{\infty} \int_{x_2-h+\gamma}^{\infty} e^{-\gamma(x_3-x_2)} dx_3 dx_2 \times \det \begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h+\gamma) & \Phi(-x_3-2h+\gamma) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h+\gamma) & \Phi(-x-x_3-h+\gamma) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3+\gamma) & \Phi(x_2-x_3+\gamma) \\ \varphi(x_3+3h-\gamma+x) & \varphi(x_3+2h-\gamma-x_2) & \varphi(h) & \Phi(h) \end{bmatrix}.$$

To compute the approximation (1.3.31) one needs to numerically evaluate a two-dimensional integral which is a routine problem for modern computers.

Correcting approximation (1.3.31) for discrete time can be performed in the same manner as correcting the ARL approximations in Section 1.3.2.1. This results in the approximation

$$\mathcal{P}_\xi(h, A, L) \cong 1 - \frac{F_{h_L, 0, -\gamma, \gamma}(3|0)}{F_{h_L, 0}(1|0)}, \text{ where } h_L := h + \omega_L, \omega_L := \sqrt{2\rho}/\sqrt{L}. \quad (1.3.32)$$

The quantity ω_L corresponds to the specialised discrete time correction of D. Siegmund (see Section 2.4.5.2 in Chapter 2 and/or Section 5.4.3.2 in Chapter 5).

In Figures 1.7, the thicker black dashed line corresponds to the empirical values of the boundary crossing probability $\mathcal{P}_\xi(h, A, L)$ computed from 100,000 simulations with different values of L and γ , where $\mu = 0$. The solid red line corresponds to the approximation in (1.3.32). The dot-dashed blue line corresponds to the diffusion approximation given in (1.3.31). The axis are: the x -axis shows the value of γ . The y -axis denotes the probabilities of reaching the barrier. The graphs, therefore, show the empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ and values of approximation (1.3.32).

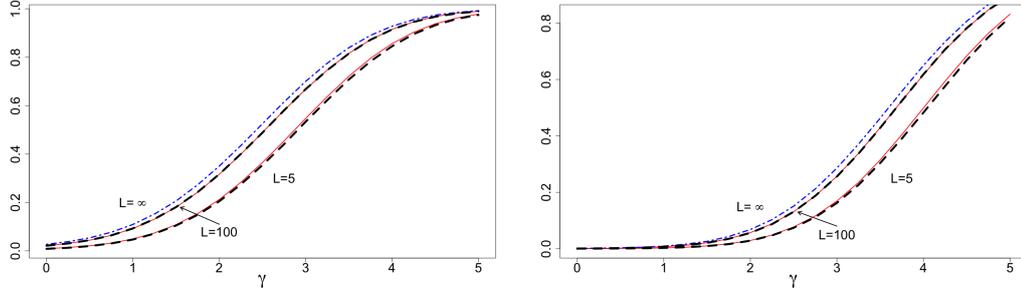


Figure 1.7: Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (thick dashed black) and its Approximation 1 (solid red) for two different values of h . Left: $h = 3$. Right: $h = 4$.

From Figure 1.7, we see that approximation (1.3.32) is very accurate even for a very small $L = 5$. We also see the significance of the discrete-time correction; whilst the diffusion approximation provides sensible results should you compare it with $L = 100$, for $L = 5$ the diffusion approximation is very far off.

1.3.4 Comparison of tests

In this section, we compare the power of the MOSUM test in (1.3.10) against the generalised MOSUM statistic (1.3.19) and the CUSUM test given in (1.2.2) specialised for this Gaussian example when used to detect a transient change. Further comparisons will be made in Section 5.5 of Chapter 5. The CUSUM statistic in (1.2.1) can be expressed as:

$$V_n = \max_{0 \leq \nu \leq n-1} \prod_{j=\nu+1}^n \exp\left(\frac{(y_j - \mu)^2 - (y_j - \mu - A)^2}{2}\right),$$

with the CUSUM test being

$$\tau_V(H) = \inf\{n \geq 1 : V_n > H\}.$$

(The choice of H will be discussed shortly.) Secondly, but also simultaneously, we compare the power of the MOSUM test as $\lambda = l/L$ varies in $[0.5, 2]$; the purpose is to demonstrate when the generalised MOSUM statistic becomes beneficial when the exact value of l is unknown and we make a potentially poor guess in the MOSUM test. This corresponds to a reasonable choice of $l/l_0 = 2$ and $l/l_1 = 0.5$. Here, we shall consider the power criterion given in (1.3.27) and set $T = 2l$. That is, we want to detect the presence of the change point within $2l - 1$ after its occurrence. For the MOSUM test, the power is then

$$\mathcal{P}_S(H_1, A, L) := \lim_{\nu \rightarrow \infty} \Pr_{\nu} \{S_{n,L} > H_1 \text{ for some } n \in [\nu - L + 1, \nu - L + 2l - 1] \mid \tau_{S,L}(H_1) > \nu - L\}.$$

For the generalised MOSUM test, the power is

$$\mathcal{P}_Z(H_2, A, l_0, l_1) := \lim_{\nu \rightarrow \infty} \Pr_{\nu} \{Z_n(l_0, l_1) > H_2 \text{ for some } n \in [\nu + 1, \nu + 2l - 1] \mid \tau_Z(H_2) > \nu\}.$$

The power of the CUSUM test for the transient change considered in then equivalent to

$$\mathcal{P}_V(H_3, A) := \lim_{\nu \rightarrow \infty} \Pr_{\nu} \{V_n > H_3 \text{ for some } n \in [\nu + 1, \nu + 2l - 1] \mid \tau_V(H_3) > \nu\}.$$

To compare the three tests, the thresholds H_1 , H_2 and H_3 have been set such that $\mathbb{E}_{\infty} \tau_M(H_1) = \mathbb{E}_{\infty} \tau_Z(H_2) = \mathbb{E}_{\infty} \tau_V(H_3) = 500$. Determination of H_1 for MOSUM has been computed using the accurate approximation in (1.3.17). For the generalised MOSUM procedure, H_2 is found via Monte Carlo simulations with 50,000 repetitions. Determination of H_3 for CUSUM was obtained using tabulated values given in [77, p. 3237].

In the first example shown in Figure 1.8 (left), we have set $A = 1$ and $l = 10$. For the MOSUM test, we consider values of $L \in [5, 20]$ to ensure $\lambda \in [0.5, 2]$. For each λ , the values of $\mathcal{P}_S(H_1, A, L)$ can be accurately approximated using the results of Chapter 5 or via Monte Carlo methods and are displayed with a solid black line. The dashed orange line depicts $\mathcal{P}_Z(H_2, A, 5, 20)$ which corresponds to prior knowledge that l is between $[5, 20]$. The shorter dashed blue line corresponds to $\mathcal{P}_V(H_3, A)$ which has been obtained via Monte Carlo simulations. In Figure 1.8 (right), we set $A = 0.5$ and $l = 20$. For the MOSUM procedure, we consider values of $L \in [10, 40]$ to ensure $\lambda \in [0.5, 2]$. In this figure, the dashed orange line depicts $\mathcal{P}_Z(H_2, A, 10, 40)$ which corresponds to prior knowledge that l is between $[10, 40]$. The shorter dashed blue line corresponds to $\mathcal{P}_V(H_3, A)$ obtained via Monte Carlo simulations. In all Monte Carlo simulations, we have used 50,000 repetitions.

From Figure 1.8 (left) and (right), one can observe the advantage of knowing l since the largest value of $\mathcal{P}_S(H_1, A, L)$ is the largest power of all three tests and is obtained for $\lambda = l/L = 1$. In these figures, the values of $\lambda = l/L$ such that $\mathcal{P}_S(H_1, A, L)$ exceeds $\mathcal{P}_Z(H_2, A, 5, 20)$ (left) and $\mathcal{P}_Z(H_2, A, 10, 40)$ (right) shows the freedom in the choice of L such that when l is unknown, you still benefit over only assuming l is bounded (similarly for CUSUM case when considering the dashed blue

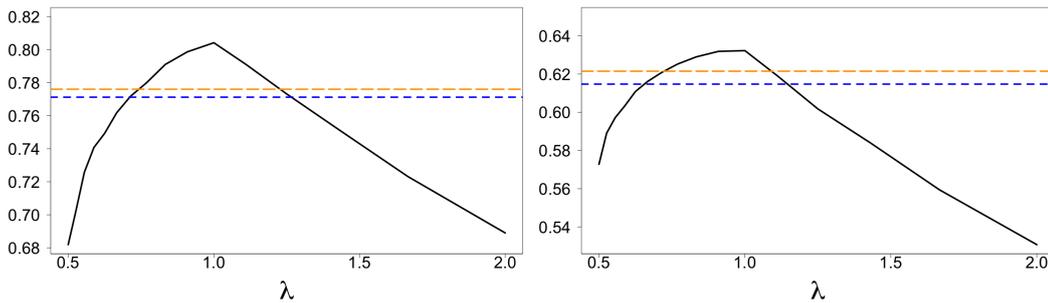


Figure 1.8: Left: Power of three tests with $A = 1$ and $l = 10$ and $ARL = 500$. Right: Power of three tests with $A = 0.5$ and $l = 20$ and $ARL = 500$.

line). From these figures it is clear that unless you are very fortunate in choosing L close to l for the MOSUM test, one should use the generalised MOSUM test if A is known. Unfortunately, there are no convenient analytic results for this test. Moreover, both the generalised MOSUM procedure and CUSUM procedures require the additional knowledge of A ; this is not true for MOSUM. For the choice of parameters considered in both examples, the additional knowledge of a transient change leads to obvious benefits in power; those is seen by comparing the generalised MOSUM orange lines with the blue CUSUM lines. Of course, $\mathcal{P}_Z(H_2, A, l_0, l_1) \rightarrow \mathcal{P}_S(H_1, A, l)$ as $l_0, l_1 \rightarrow l$.

1.3.5 An application to real world data

Hydrostatic pressure testing is important safety precaution for the Oil and Gas industry, see [70]. Pressure testing is performed to confirm a pressure containing system is structurally sound and not leaking. Tests are performed by increasing the pressure in the system, expanding the pressure body, until the pressure reaches a pre-defined value typically equal to or larger than the body rated design pressure, then holding it there for a long enough time period to confirm there are no leaks, until eventually releasing the pressure. When performing tests offshore on floating Vessel/Drilling Rigs (Rig) this is complicated by the Rig's movement due to the ocean waves, which introduce nearly sinusoidal fluctuations in pressure. Many of these tests are performed in real time and in parallel. Locating automatically when a test has been performed is essential for pressure analysis to determine if a leak is present and this is not obvious when noise is large. Typical example data is shown in Figure 1.9. Note that the data motivating this section is confidential and cannot be disclosed here. As a result, all tests that are performed are fictitious. When performing pressure tests, the hold periods can differ in length and amplitudes (pressure).

A sensible way of modelling the data under the null hypothesis of no pressure test could be $z_t = s_t + y_t$, where s_t represents the signal introduced by the wave motion and y_t can be modelled as i.i.d. $N(\mu, \sigma^2)$ and reflects the random noise that is present in the system. In most scenarios, there is significant pre-test data so s_t, μ and σ can be estimated with great accuracy and therefore assumed known. How to estimate s_t or in general how to remove all main components of a signal leaving

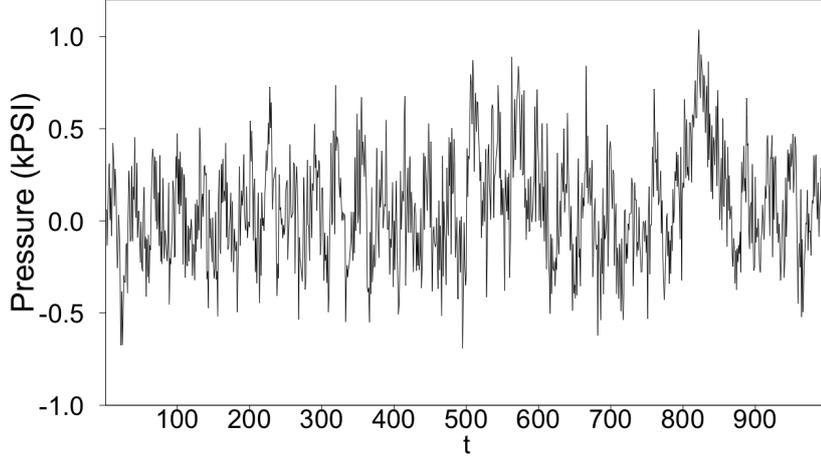


Figure 1.9: Typical pressure data

only noise can be performed using Singular Spectrum Analysis, see [35, 36]. When a pressure test begins, this can be reflected with a change in mean of the y_t ; that is, under a pressure test $\mathbb{E}y_t = \mu + A$. The value of A is often constant, but can differ between tests and is generally unknown. Each test can differ in duration but typical lengths vary between $l \in [50, 100]$ units of time. One has to detect a transient change in mean of $y_t = z_t - s_t$. The behaviour of $z_t - s_t$ is shown in Figure 1.10 (left). In Figure 1.10 (right), we depict the MOSUM statistic setting $L = 75$. The horizontal line in this figure corresponds to the threshold required for an ARL of 5000. Note that the choice of ARL depends on the users tolerance to false alarms and has been fixed at 5000 as an example. The MOSUM statistic indicates the location of three performed pressure tests and has the great advantage of not requiring knowledge A when determining the ARL threshold unlike the generalised MOSUM and CUSUM procedures. A similar example is shown in Figure 1.11, where $L = 150$ has been selected; three tests have been clearly located. Note that in Figure 1.10 (right) and Figure 1.11 (right), the MOSUM statistic is depicted with a shift in time by L ($t \rightarrow t - L$). This explains the early exceedance seen in Figure 1.11 (right).

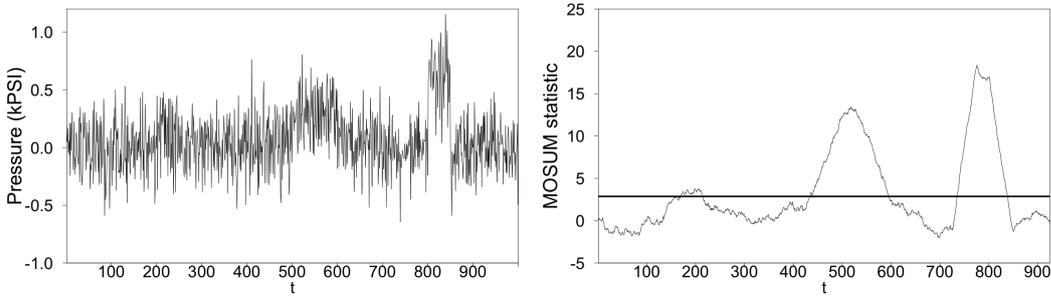


Figure 1.10: Left: Behaviour of y_t . Right: MOSUM statistic with $L = 50$ and ARL= 5000.

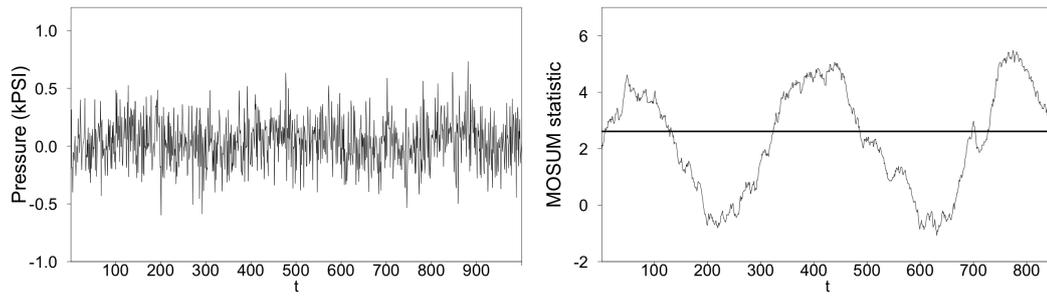


Figure 1.11: Left: Behaviour of y_t . Right: MOSUM statistic with $L = 150$ and $ARL= 5000$.

Chapter 2

Approximations for the boundary crossing probabilities of moving sums of normal random variables

Abstract

In this chapter, we study approximations for boundary crossing probabilities for the moving sums of i.i.d. normal random variables. As will become clear later, the boundary crossing probabilities studied in this chapter are strongly related to the MOSUM test discussed in Chapter 1. To develop our approximations, we approximate a discrete time problem with a continuous time problem, allowing the application of developed theory for stationary Gaussian processes and a number of approximations (some well known and some not). Particular attention is paid to the strong performance of a newly developed approximation that corrects the use of continuous time results in a discrete time setting. Results of extensive numerical comparisons are reported. These results show that the developed approximations are very accurate even for small window length. Also, they have high accuracy when the original r.v. are not exactly normal and when the weights in the moving window are not all equal. Accurate and simple approximations are then provided for ARL, the average run length until crossing the boundary. The content of this chapter contains results from the two published papers [82, 83].

2.1 Introduction: Statement of the problem

Let $\varepsilon_1, \varepsilon_2, \dots$ be a sequence of i.i.d. normal random variables (r.v.) with mean μ and variance $\sigma^2 > 0$. For a fixed positive integer L , the moving sums are defined by

$$S_{n,L} := \sum_{j=n+1}^{n+L} \varepsilon_j \quad (n = 0, 1, \dots). \quad (2.1.1)$$

The sequence of the moving sums (2.1.1) will be denoted by \mathbb{S} so that $\mathbb{S} = \{S_{0,L}, S_{1,L}, \dots\}$.

The main aim of this chapter is to develop accurate approximations for the following related characteristics of \mathbb{S} (note that for the sake of simplicity of notation we are not indicating the dependence of these characteristics on L).

- (a) The boundary crossing probability (BCP) for the maximum of the moving sums:

$$\mathcal{P}_{\mathbb{S}}(M, H) := \Pr \left(\max_{n=0,1,\dots,M} S_{n,L} \geq H \right), \quad (2.1.2)$$

where M is a given positive integer and H is a fixed threshold. Note that the total number of r.v. ε_i used in (2.1.2) is $M+L$ and $\mathcal{P}_{\mathbb{S}}(M, H) \rightarrow 1$ as $M \rightarrow \infty$, for all H and L . We aim to provide an approximation to $\mathcal{P}_{\mathbb{S}}(M, H)$ that is uniformly accurate for all H and not just for large H .

- (b) The probability distribution of the moment of time $\tau_H(\mathbb{S}) := \min\{n \geq 0 : S_{n,L} \geq H\}$ when the sequence $S_{n,L}$ reaches the threshold H for the first time. The BCP $\mathcal{P}_{\mathbb{S}}(M, H)$, considered as a function of M , is the c.d.f. of this probability distribution: $\mathcal{P}_{\mathbb{S}}(M, H) = \Pr(\tau_H(\mathbb{S}) \leq M)$.

- (c) The average run length (ARL) until \mathbb{S} reaches H for the first time:

$$\text{ARL}_H(\mathbb{S}) := \sum_{n=0}^{\infty} n \Pr\{\tau_H = n\} = \int_0^{\infty} M d\mathcal{P}_{\mathbb{S}}(M, H). \quad (2.1.3)$$

Developing accurate approximations for the BCP $\mathcal{P}_{\mathbb{S}}(M, H)$ and the associated ARL (2.1.3) for generic parameters H , M , L is very important in various areas of statistics, predominantly in applications related to change-point detection as is discussed in Chapter 1; see also [18, 33, 34, 74]. Engineering applications of MOSUM (moving sums charts) are extremely important and have been widely discussed in literature; see e.g. [5, 32, 33, 136]. The BCP $\mathcal{P}_{\mathbb{S}}(M, H)$ is an $(M+1)$ -dimensional integral and therefore direct evaluation of this BCP is hardly possible even with modern software.

To derive approximations for the BCP (2.1.2) one can use some generic approximations such as Durbin and Poisson Clumping Heuristic considered below. These approximations, however, are not accurate especially for small window length L ; this is demonstrated below in this chapter. Furthermore, they only begin to work for large H . There is, therefore, a need for derivation of specific approximations for the BCP (2.1.2) and the ARL (2.1.3). Such a need was well understood in the statistical community and indeed very accurate approximations for the BCP and the ARL have been developed in a series of papers by J. Glaz and coauthors, see for example [30, 34, 137, 138] (the methodology was also extended to the case when ε_j are integer-valued r.v., see [31]). We will call these approximations 'Glaz approximations' by the name of the main author of these papers; they will be formally written down in Sections 2.4 and 2.8.

The approximations developed in this chapter take two fundamentally different forms depending on whether $M \leq L$ or $M > L$. To derive the approximations when $M \leq L$, we will use the methodology developed in [143, Ch.2,§2] for the continuous-time case, which has to be modified for discrete time. The approximations developed for $M \leq L$ are able to utilise the conditionally Gauss-Markov property which is present for the continuous time analogue of \mathbb{S} . For $M > L$, we will formulate a number of approximations by correcting the continuous time results of L. Shepp

in [113] for discreteness. The accuracy of these approximations is very high and similar to the Glaz approximations; this is discussed in Sections 2.6 and 2.8. The methodologies of derivation of Glaz approximations and the approximations of this chapter are very different. The practical advantage of the approximations developed in this chapter (they require approximating either a one-dimensional integral or an eigenvalue of an integral operator) is their relative simplicity as to compute the Glaz approximations one needs to numerically approximate $L + 1$ and $2L + 1$ dimensional integrals. This is not an easy task even taking into account the fact of existence of a sophisticated software; see references in Section 2.4.1. This chapter is structured as follows.

In Section 2.2 we reformulate the problem and discuss how to approximate our discrete-time problem with a continuous-time problem. Here we state a number of classical and simple approximations. In Section 2.3, we consider the case of $M \leq L$ and derive new approximations that correct the use of continuous time results in a discrete time setting; they will be referred to as the ‘Corrected Diffusion Approximations’ or simply CDA’s. In Section 2.4, we consider the case when $M > L$ and begin by stating the Glaz approximation. We subsequently provide exact formulas for the first-passage probabilities (in the continuous-time setup) due to L. Shepp [113] and give their alternative representation which will be crucial for deriving some of our approximations. In this section we adapt the methodology of D. Siegmund to correct Shepp’s formulas for discrete time and define a version of the Glaz approximation which we will call Glaz-Shepp-Siegmund approximation. In Section 2.5 we develop continuous-time approximations based on approximating eigenvalues of integral operators and subsequently correct them for discrete time. In Section 2.6 we present results of large-scale simulation studies evaluating the performance of the approximations (also, in the cases when the original r.v. ε_j are not normal and the weights in the moving window are not equal). In Section 2.7, we briefly discuss an approximation with connections to Extreme value theory and assess its accuracy. In Section 2.8, we develop an approximation for $\text{ARL}_H(\mathbb{S})$ and compare its accuracy to the one developed in [34].

2.2 Boundary crossing probabilities and related characteristics: discrete and continuous time

2.2.1 Standardisation of the moving sums

For convenience, we standardise the moving sums $S_{n,L}$ defined in (2.1.1).

The first two moments of $S_{n,L}$ are

$$\mathbb{E} S_{n,L} = \mu L, \quad \text{var}(S_{n,L}) = \sigma^2 L. \quad (2.2.1)$$

Define the standardized r.v.’s:

$$\xi_{n,L} := \frac{S_{n,L} - \mathbb{E} S_{n,L}}{\sqrt{\text{var}(S_{n,L})}} = \frac{S_{n,L} - \mu L}{\sigma \sqrt{L}}, \quad n = 0, 1, \dots, \quad (2.2.2)$$

and denote $\mathbb{X} = \{\xi_{0,L}, \xi_{1,L}, \dots\}$. All r.v. $\xi_{n,L}$ are $N(0, 1)$; that is, they have the probability density function and c.d.f.

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \Phi(t) := \int_{-\infty}^t \varphi(x) dx. \quad (2.2.3)$$

Unlike the original r.v. ε_i , the r.v. $\xi_{0,L}, \xi_{1,L}, \dots$ are correlated with correlations depending on L , see Section 2.2.2 below.

The BCP $\mathcal{P}_{\mathbb{S}}(M, H)$ defined by (2.1.2) is equal to the BCP

$$\mathcal{P}_{\mathbb{X}}(M, h) := \Pr \left(\max_{n=0,1,\dots,M} \xi_{n,L} \geq h \right), \quad (2.2.4)$$

where

$$H = \mu L + \sigma h \sqrt{L} \quad \text{so that} \quad h = \frac{H - \mu L}{\sigma \sqrt{L}}. \quad (2.2.5)$$

Similarly, $\tau_H(\mathbb{S}) = \tau_h(\mathbb{X})$ and $\text{ARL}_H(\mathbb{S}) = \text{ARL}_h(\mathbb{X})$.

Note that studying the probability $\mathcal{P}_{\mathbb{X}}(M, h)$ is equivalent to studying the probability

$$F_{\mathbb{X}}(M, h) := \Pr \left(\max_{n=0,1,\dots,M} \xi_{n,L} < h \right), \quad (2.2.6)$$

where $\mathcal{P}_{\mathbb{X}}(M, h) = 1 - F_{\mathbb{X}}(M, h)$. In accordance with the terminology of [122] and [113] we shall call $F_{\mathbb{X}}(M, h)$ ‘first-passage probability’.

In what follows, we derive approximations for (2.2.4) and hence the distribution of $\tau_h(\mathbb{X})$ and $\text{ARL}_h(\mathbb{X})$. These approximations will be based on approximating the sequence $\{\xi_{i,L}\}_i$ by a continuous time random process and subsequently correcting the obtained approximations for discreteness.

2.2.2 Correlation between $\xi_{n,L}$ and $\xi_{n+k,L}$

In order to derive our approximations, we will need explicit expressions for the correlations $\text{Corr}(\xi_{n,L}, \xi_{n+k,L})$. For a proof of the following Lemma, see Appendix A in Section 2.9.

Lemma 2.2.1 *Let $\xi_{n,L}$ be as defined in (2.2.2). Then $\text{Corr}(\xi_{0,L}, \xi_{k,L}) = \text{Corr}(\xi_{n,L}, \xi_{n+k,L})$ and*

$$\text{Corr}(\xi_{0,L}, \xi_{k,L}) = \frac{\mathbb{E}(\xi_{0,L} \xi_{k,L}) - (\mathbb{E} \xi_{0,L})^2}{\text{var}(\xi_{0,L})} = 1 - k/L. \quad (2.2.7)$$

for $0 \leq k \leq L$. If $k > L$ then $\text{Corr}(\xi_{0,L}, \xi_{k,L}) = 0$.

2.2.3 Continuous-time (diffusion) approximation

For the purpose of approximating the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ and the associated characteristics introduced in Introduction, we replace the discrete-time process $\xi_{0,L}, \dots, \xi_{M,L}$ with a continuous process $S(t)$, $t \in [0, T]$, where $T = M/L$. This is done as follows.

Set $\Delta = 1/L$ and define $t_n = n\Delta \in [0, T]$ $n = 0, 1, \dots, M$. Define a piece-wise linear continuous-time process $S_L(t)$, $t \in [0, T]$:

$$S_L(t) = \frac{1}{\Delta} [(t_n - t)\xi_{n-1,L} + (t - t_{n-1})\xi_{n,L}] \quad \text{for } t \in [t_{n-1}, t_n], \quad n = 1, \dots, M.$$

By construction, the process $S_L(t)$ is such that $S_L(t_n) = \xi_{n,L}$ for $n = 0, \dots, M$. Also we have that $S_L(t)$ is a second-order stationary process in the sense that $\mathbb{E} S_L(t)$, $\text{var}(S_L(t))$ and the autocorrelation function $R^{(L)}(t, t+k\Delta) = \text{Corr}(S_L(t), S_L(t+k\Delta))$ do not depend on t .

Lemma 2.2.2 *Assume $L \rightarrow \infty$. The limiting process $S(t) = \lim_{L \rightarrow \infty} S_L(t)$, where $t \in [0, T]$, is a Gaussian second-order stationary process with marginal distribution $S(t) \sim N(0, 1)$ for all $t \in [0, T]$ and autocorrelation function $R(t, t+s) = R(s) = \max\{0, 1-|s|\}$.*

This lemma is a simple consequence of Lemma 2.2.1. Note that although we are letting L tend to infinity (also $M \rightarrow \infty$ since T is fixed), subsequent results will be corrected to produce approximations for finite L .

2.2.4 Diffusion approximations for the main characteristics of interest

The above approximation of a discrete-time process \mathbb{S} with a continuous process $S(t)$, $t \in [0, T]$, allows us to approximate the characteristics introduced in Introduction by the continuous-time analogues as follows.

- (a) BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ is approximated by $P(T, h)$, which is the probability of reaching the threshold h by the process $S(t)$ on the interval $[0, T]$:

$$P(T, h) := \Pr \left\{ \max_{0 \leq t \leq T} S(t) \geq h \right\} = \Pr \left\{ S(t) \geq h \text{ for at least one } t \in [0, T] \right\}. \quad (2.2.8)$$

Note that $P(0, h) = 1 - \Phi(h) > 0$.

At times, it will be convenient to use the first-passage probability

$$F(T, h) = \Pr \left\{ \max_{0 \leq t \leq T} S(t) < h \right\} = 1 - P(T, h).$$

Since $\xi_{0,L} = S(0) \sim N(0, 1)$, we have $F(0, h) = 1 - P(0, h) = \Phi(h)$.

- (b) The time moment $\tau_H(\mathbb{S}) = \tau_h(\mathbb{X})$ is approximated by $\tau_h(S(t)) := \min\{t \geq 0 : S(t) \geq h\}$, which is the time moment when the process $S(t)$ reaches h . The distribution of $\tau_h(S(t))$ has the form:

$$(1 - \Phi(h))\delta_0(ds) + q(s, h, S(t))ds, \quad s \geq 0,$$

where $\delta_0(ds)$ is the delta-measure concentrated at 0 and

$$q(s, h, S(t)) = \frac{d}{ds} P(s, h), \quad 0 < s < \infty. \quad (2.2.9)$$

The function $q(s, h, S(t))/\Phi(h)$, considered as a function of s , is a probability density function on $(0, \infty)$ since

$$\int_0^\infty q(s, h, S(t)) ds = 1 - P(0, h) = \Phi(h).$$

(c) $\text{ARL}_h(\mathbb{X})/L$ is approximated by

$$\text{ARL}_h(S(t)) = \mathbb{E}[\tau_h(S(t))] = \int_0^\infty s q(s, h, S(t)) ds. \quad (2.2.10)$$

We will call approximations (2.2.8) and (2.2.10) diffusion approximations, see Section 2.3.1. As indicated by numerical results shown in Section 2.3.4, if L and M are very large then the diffusion approximations are rather accurate. For not very large values of L and M these approximations will be much improved with the help of the methodology developed by D.Siegmund and adapted to our setup in Sections 2.3.3 and 2.4.5.

2.2.5 Durbin and Poisson Clumping approximations for the BCP $P(T, h)$

Derivation of the exact formulas for the BCP $P(T, h)$ has been discussed in several papers including [71, 112, 113, 114, 143]; exact formulas will be provided in Sections 2.3.1 and 2.4.2.

In this section, we provide explicit formulas for two simple approximations for the BCP $P(T, h)$ based on general principles. We will assess the accuracy of these approximations in Section 2.3.4 and will find that the accuracy of both of them is quite poor. The purpose of including these two approximations into our collection is only to demonstrate that the original problems stated in Introduction are not easy and cannot be handled by general-purpose techniques. More sophisticated techniques using the specificity of the problem should be used, which is exactly what is done in this chapter. The first generic approximation considered is the Durbin approximation which is constructed on the base of [22] and is explained in Appendix B in Section 2.10.

Approximation 1. *Durbin approximation for the BCP (2.2.8): $P(T, h) \cong hT \varphi(h)$.*

Let us now state the second approximation for the BCP defined in (2.2.8), which is the Poisson Clumping Heuristic (PCH) formulated as Lemma 2.2.3 according to [2] p. 81; the PCH heuristic appears in many places within the boundary crossing literature, see [120, 121].

Lemma 2.2.3 *Let $X(t)$ be a stationary Gaussian process with mean zero and covariance function satisfying $R(t) = 1 - |t|$ as $t \rightarrow 0$. Then for large h , $T_h = \min\{t : X(t) \geq h\}$ is approximately exponential with parameter $h\varphi(h)$.*

From Lemma 2.2.3 we obtain:

Approximation 2. *PCH approximation for BCP (2.2.8):*

$$P(T, h) \cong 1 - \exp(-h\varphi(h)T).$$

As can be seen from Fig. 2.1 and Fig. 2.2 in Section 2.3.4, Approximations 1 and 2 are poor approximations for $P(T, h)$ and $\mathcal{P}_{\mathbb{X}}(M, h)$ when $M/L \leq 1$ and $M/L > 1$; the case $M/L > 1$ is discussed in Section 4.6 in [82] and is not included in this thesis.

2.3 Diffusion approximation with and without discrete-time correction; $M \leq L$

In this section, we assume $M \leq L$ and hence $T = M/L \leq 1$. The more complicated case $M > L$ will be considered in Section 2.4 and beyond.

2.3.1 Diffusion approximation, formulation

Here we collect explicit formulas for the BCP $P(T, h)$ that can be obtained in [71, 122, 143]; the proofs are given in Section 2.3.2 to provide insight for later results. We have:

$$P(T, h) = 1 - \Phi^2(h) + \varphi(h)[h\Phi(h) + \varphi(h)], \quad T = 1; \quad (2.3.1)$$

$$\begin{aligned} P(T, h) = & 1 - \int_{-\infty}^h \Phi\left(\frac{h(Z+1)-x(-Z+1)}{2\sqrt{Z}}\right) \varphi(x) dx + \\ & + \frac{2\sqrt{Z}}{Z+1} \varphi(h) \left[h\sqrt{Z} \Phi(h\sqrt{Z}) + \frac{1}{\sqrt{2\pi}} (\sqrt{2\pi}\varphi(h))^Z \right], \quad 0 < T \leq 1, \end{aligned} \quad (2.3.2)$$

where $Z = T/(2 - T)$. If $T = 1$ then (2.3.2) simplifies to (2.3.1). We refer to the above stated formulas for $P(T, h)$ as Approximation 3 or ‘Diffusion approximation’.

Approximation 3. *The Diffusion approximation for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ defined in (2.2.4) in case $M \leq L$: formula (2.3.2) with $T = M/L$; if $M = L$ then (2.3.2) reduces to (2.3.1).*

In Section 2.3.3, we will derive a discrete-time correction for the Diffusion approximation. In order to do this, we need to correct the steps used for deriving (2.3.2). This explains that, despite the formula (2.3.2) is known, we need to derive it (in order to correct certain steps of its derivation). This is done in the next section which follows from [143], p.69.

2.3.2 Derivation of (2.3.2)

2.3.2.1 Conditioning on the initial value.

From Lemma 2.2.2, $\{S(t), t \in [0, \infty)\}$, is a stationary Gaussian process with mean $\mathbb{E} S(t) = 0$ and covariance function $\mathbb{E} S(t)S(t+u) = \max\{0, 1 - |u|\}$. By conditioning on the initial state of the process $S(t)$, we define

$$Q_h(T, x_0) := \Pr \left\{ \max_{t \in [0, T]} S(t) > h \mid S(0) = x_0 \right\}.$$

Since $x_0 \sim N(0, 1)$ the BCP $P(T, h)$ is

$$P(T, h) = \int_{-\infty}^h Q_h(T, x_0) \varphi(x_0) dx_0 + 1 - \Phi(h), \quad (2.3.3)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ are defined in (2.2.3). In order to proceed we seek an explicit expression for $Q_h(T, x_0)$. We shall firstly discuss a known BCP formula for the Brownian motion before returning to explicit evaluation of $Q_h(T, x_0)$.

2.3.2.2 Boundary crossing probabilities for the Brownian Motion.

Let $W(t)$ be the standard Brownian Motion process on $[0, \infty)$ with $W(0) = 0$ and $\mathbb{E}W(t)W(s) = \min(t, s)$. For given $a, R > 0$ and $b \in \mathbb{R}$, define

$$P_W(R; a, b) := \Pr \{W(t) > a + bt \text{ for at least one } t \in [0, R]\}, \quad (2.3.4)$$

which is the probability that the Brownian motion $W(t)$ reaches a sloped boundary $a + bt$ within the time interval $[0, R]$. Using results of [118], for any $a, R > 0$ and any real b we have

$$P_W(R; a, b) = 1 - \Phi\left(\frac{bR + a}{\sqrt{R}}\right) + e^{-2ab} \Phi\left(\frac{bR - a}{\sqrt{R}}\right). \quad (2.3.5)$$

In particular, for $R = 1$ we have

$$P_W(1; a, b) = 1 - \Phi(b + a) + e^{-2ab} \Phi(b - a). \quad (2.3.6)$$

2.3.2.3 Boundary crossing probabilities for $S(t)$.

Let $\{S_0(t), t \in [0, \infty)\}$ be a process obtained by considering only the sample functions of $\{S(t), t \in [0, \infty)\}$ which are equal to x_0 at $t = 0$. For $0 \leq t \leq 1$, we obtain from [71], p.520, that $S_0(t)$ can be expressed in terms of the Brownian motion:

$$S_0(t) = (2 - t)W(g(t)) + x_0(1 - t) \quad (2.3.7)$$

with $g(t) = t/(2 - t)$. It then follows from (2.3.7) that for $T \leq 1$ and $x_0 < h$ we have

$$\begin{aligned} Q_h(T, x_0) &= \Pr\{S_0(t) \geq h \text{ for at least one } t \in [0, T]\} \\ &= \Pr\left\{W(g(t)) \geq \frac{h - x_0(1 - t)}{2 - t} \text{ for at least one } t \in [0, T]\right\}. \end{aligned}$$

Noting that $t = 2g(t)/(1 + g(t))$ we obtain

$$\begin{aligned} Q_h(T, x_0) &= \Pr\left\{W(g(t)) \geq \left(\frac{(h - x_0)(1 + g(t))}{2}\right) + x_0g(t) \text{ for at least one } t \in [0, T]\right\} \\ &= \Pr\left\{W(t') \geq \left(\frac{h - x_0}{2}\right) + t'\left(\frac{h + x_0}{2}\right) \text{ for at least one } t' \in \left[0, \frac{T}{2 - T}\right]\right\} \\ &= P_W(Z; a, b), \end{aligned} \quad (2.3.8)$$

where $Z = T/(2 - T)$, $b = (h + x_0)/2$ and $a = (h - x_0)/2$. Using (2.3.5), we conclude

$$Q_h(T, x_0) = 1 - \Phi\left(\frac{bZ + a}{\sqrt{Z}}\right) + e^{-2ab} \Phi\left(\frac{bZ - a}{\sqrt{Z}}\right).$$

One can then show that by using this explicit form for $Q_h(T, x_0)$ in the integral (2.3.3), we obtain (2.3.1) and (2.3.2).

It is now evident how BCP formula (2.3.5) for the Brownian motion can be used to obtain (2.3.1) and (2.3.2). To improve the diffusion approximations for discrete time, we aim at correcting the conditional probability $Q_h(T, x_0)$ for discrete time. Because of the relation shown in (2.3.8), the approach taken in this chapter is to correct (2.3.5) for discrete time.

2.3.3 Discrete Time Correction

2.3.3.1 Discrete time correction for the BCP of cumulative sums.

Let X_1, X_2, \dots be i.i.d. $N(0, 1)$ r.v.'s and set $Y_n = X_1 + X_2 + \dots + X_n$. Consider the sequence of cumulative sums $\{Y_n\}$ and define the stopping time $\tau_{Y,a,b} = \inf\{n \geq 1 : Y_n \geq a + bn\}$ for $a > 0$ and $b \in \mathbb{R}$. Consider the problem of evaluating

$$\Pr(\tau_{Y,a,b} \leq N) = \Pr(Y_n \geq a + bn \text{ for at least one } n \in \{1, 2, \dots, N\}). \quad (2.3.9)$$

Exact evaluation of (2.3.9) is difficult even if N is not very large but it was accurately approximated by D. Siegmund see e.g. [118, p. 19]. Let $W(t)$ be the standard Brownian Motion process on $[0, \infty)$. For $a > 0$ and $b \in \mathbb{R}$, define $\tau_{W,a,b} = \inf\{t : W(t) \geq a + bt\}$ so that

$$\Pr(\tau_{W,a,b} \leq N) = P_W(N, a + bt). \quad (2.3.10)$$

In [118], (2.3.10) was used to approximate (2.3.9) after translating the barrier $a + bt$ by a suitable scalar $\rho \geq 0$. Specifically, the following approximation has been constructed:

$$P(\tau_{Y,a,b} \leq N) \cong P_W(N, (a + \rho) + bt),$$

where the constant ρ approximates the expected excess of the process $\{Y_n\}$ over the barrier $a + bt$. From [117] (p. 225)

$$\rho = -\pi^{-1} \int_0^\infty \lambda^{-2} \log\{2(1 - \exp(-\lambda^2/2))/\lambda^2\} d\lambda \cong 0.5826. \quad (2.3.11)$$

Whence, by denoting $\hat{a} = a + \rho$ and recalling (2.3.5), D. Siegmund's formulas of [118] imply the approximation:

$$\Pr(\tau_{Y,a,b} \leq N) \cong \Pr(\tau_{W,\hat{a},b} \leq N) = 1 - \Phi\left(\frac{bN + \hat{a}}{\sqrt{N}}\right) + e^{-2ab} \Phi\left(\frac{bN - \hat{a}}{\sqrt{N}}\right).$$

2.3.3.2 Discretized Brownian motion.

In this section, we modify D. Siegmund arguments discussed in previous section to the case when the r.v. are indexed by points on the uniform grid in an interval and therefore the sequence of cumulative sums compares with a limiting Brownian motion process which lies within this interval.

Assume that $Z > 0$ and M is a positive integer. Define $\epsilon = Z/M$ and let $t'_n = n\epsilon \in [0, Z]$, $n = 0, 1, \dots, M$. Let X_1, X_2, \dots be i.i.d. $N(0, 1)$ r.v.'s and set $W(t'_n) = \sqrt{\epsilon} \sum_{i=1}^n X_i$. For $a > 0$ and $b \in \mathbb{R}$, define the stopping time

$$\tau_{W,a,b} = \inf\{t'_n : W(t'_n) \geq a + bt'_n\} \quad (2.3.12)$$

and consider the problem of approximating

$$\Pr(\tau_{W,a,b} \leq Z) = \Pr\left(W(t'_n) \geq a + bt'_n \text{ for at least one } t'_n \in \{0, \epsilon, \dots, M\epsilon = Z\}\right). \quad (2.3.13)$$

As $M \rightarrow \infty$, the piecewise linear continuous-time process $W^\epsilon(t)$, $t \in [0, Z]$, defined by:

$$W^\epsilon(t) := \frac{1}{\epsilon} [(t'_n - t)W(t'_{n-1}) + (t - t'_{n-1})W(t'_n)] \quad \text{for } t \in [t'_{n-1}, t'_n], \quad n = 1, \dots, M,$$

converges to the Brownian motion on $[0, Z]$. For this reason, we refer to the sequence $\{W(t'_1), \dots, W(t'_M)\}$ as discretized Brownian motion. We make the following connection between $W(t'_n)$ and the random walk Y_n :

$$W(t'_n) = \sqrt{\epsilon} Y_n = Y_n / \sqrt{M/Z}, \quad n = 1, 2, \dots, M.$$

Then by using (2.3.11), we approximate the expected excess over the boundary for the process $W(t'_n)$ by $\rho_{M/Z} = 0.5826/\sqrt{M/Z}$.

Thus, using the same methodology as D. Siegmund, in order to obtain an accurate approximation for (2.3.13), we translate the barrier $a + bt$ by the discrete time correction factor $\rho_{M/Z}$ and apply (2.3.5). By denoting $\hat{a} = a + \rho_{M/Z}$, we obtain the approximation to (2.3.13):

$$\Pr(\tau_{W,a,b} \leq Z) \cong 1 - \Phi\left(\frac{bZ + \hat{a}}{\sqrt{Z}}\right) + e^{-2\hat{a}b} \Phi\left(\frac{bZ - \hat{a}}{\sqrt{Z}}\right). \quad (2.3.14)$$

2.3.3.3 Corrected Diffusion Approximation.

Let $Q_{h,\rho}(M, x_0)$ denote the discrete time corrected equivalent of $Q_h(T, x_0)$, where $T = M/L \leq 1$. Using (2.3.14) and the relation shown in (2.3.8),

$$Q_{h,\rho}(M, x_0) = 1 - \Phi\left(\frac{bZ + \hat{a}}{\sqrt{Z}}\right) + e^{-2\hat{a}b} \Phi\left(\frac{bZ - \hat{a}}{\sqrt{Z}}\right) \quad (2.3.15)$$

with

$$T = \frac{M}{L}, \quad Z = \frac{T}{2-T}, \quad \hat{a} = \frac{h-x_0}{2} + \rho_{M/Z}, \quad b = \frac{h+x_0}{2}, \quad \rho_{M/Z} = \frac{0.5826}{\sqrt{M/Z}}.$$

Using $Q_{h,\rho}(M, x_0)$ in (2.3.3), the equivalent probability $P(T, h)$ after correction for discrete time will be denoted by $P_\rho(M, h)$.

Approximation 4. For $M \leq L$ (that is, $T \leq 1$), the corrected diffusion approximation (CDA) for the BCP (2.2.4) is given by

$$\mathcal{P}_{\mathbb{X}}(M, h) \cong P_\rho(M, h) := \int_{-\infty}^h Q_{h,\rho}(M, x_0) \varphi(x_0) dx_0 + 1 - \Phi(h), \quad (2.3.16)$$

where $Q_{h,\rho}(M, x_0)$ is given in (2.3.15).

For $M = L$ we have $T = Z = 1$ and the CDA $P_\rho(M, h)$ can be explicitly evaluated:

$$P_\rho(L, h) = 1 - \Phi(h + \rho_L) \Phi(h) + \frac{\varphi(h + \rho_L)}{\rho_L} \Phi(h) - \frac{\varphi(h) e^{-2h\rho_L}}{\rho_L} \Phi(h - \rho_L), \quad (2.3.17)$$

where $\rho_L := 0.5826/\sqrt{L}$. For a proof of (2.3.17), see Appendix C in Section 2.11.

2.3.4 Simulation study, $T \leq 1$

In this section, we study the quality of the Durbin (Approximation 1), PCH (Approximation 2), Diffusion (Approximation 3) approximations and the CDA (Approximation 4) for the BCP $\mathcal{P}_\mathbb{X}(M, h)$, defined in (2.2.4), when $M \leq L$ (that is, $T \leq 1$). Without loss of generality, ε_j in (2.1.1) are normal r.v.'s with mean 0 and variance 1. In Figures 2.1–2.2, the black dashed line corresponds to the empirical values of the BCP $\mathcal{P}_\mathbb{X}(M, h)$ defined by (2.2.4) computed from 100,000 simulations with different values of L and M (for given L and M , we simulate $L + M$ normal random variables 100,000 times). For $j = 1, \dots, 4$, the number j next to a line corresponds to Approximation j . The axis are: the x -axis shows the value of the normalized barrier h , see (2.2.5); the y -axis denotes the probabilities of reaching the barrier. The graphs, therefore, show the empirical probabilities of reaching the barrier h (for the dashed line) and values of considered approximations for these probabilities.

In Table 2.1, we display the relative error of the CDA with respect to the empirical BCP $\mathcal{P}_\mathbb{X}(M, h)$ for all considered parameter choices. Numerical study of this section shows that in the case $T \leq 1$, the accuracy of the CDA (Approximation 4) is excellent, even for rather small L and M . Furthermore, the CDA is uniformly accurate in h . At the same time, the Durbin, PCH and Diffusion approximations are generally poor (note however that the accuracy of the Diffusion approximation improves as L increases). The discrete time correction factor brings a significant improvement to the Diffusion approximation resulting in a very small relative errors shown in Table 2.1.

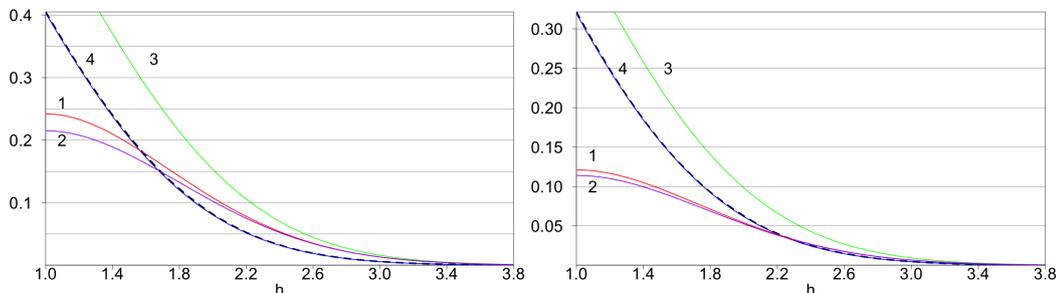


Figure 2.1: Empirical probabilities of reaching the barrier h and four approximations. Left: $L = 5$, $M = 5$, $T = 1$. Right: $L = 10$, $M = 5$, $T = 1/2$.

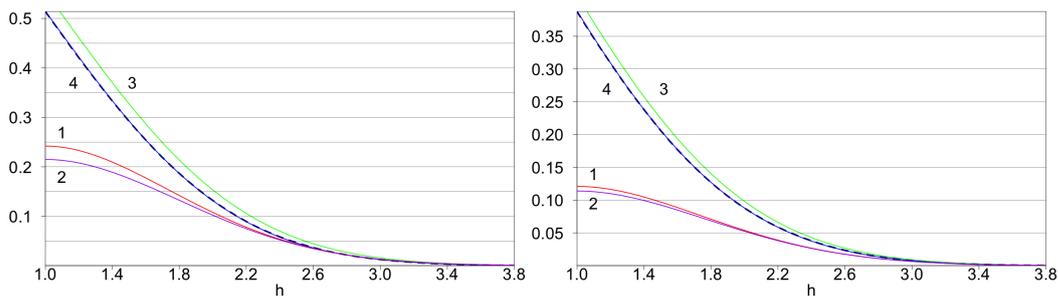


Figure 2.2: Empirical probabilities of reaching the barrier h and four approximations. Left: $L = 100$, $M = 100$, $T = 1$. Right: $L = 200$, $M = 100$, $T = 1/2$.

Table 2.1: Relative error of the CDA with respect to the empirical BCP (in percent)

BCP	$L=5, M=5$	$L=10, M=5$	$L=100, M=100$	$L=200, M=100$
0.05	0.225 %	0.238 %	0.041 %	0.132 %
0.10	0.316 %	0.284 %	0.093 %	0.103 %
0.15	0.474 %	0.326 %	0.155 %	0.059 %
0.20	0.390 %	0.296 %	0.228 %	0.101 %

2.4 Approximations for the BCP in continuous and discrete time; $M > L$

In this section, we assume M is any integer larger than L and thus $T = M/L > 1$. Apart from the obvious failure in (2.3.7) when $T > 1$, perhaps an intuitive reason why the case of $T > 1$ requires a different approach theoretically is related to correlation. On the interval $[0, T]$ with $T \leq 1$, at all values of $t, t' \in [0, T]$ we have $S(t)$ and $S(t')$ are correlated. For $T > 1$, there exists values of $t, t' \in [0, T]$ such that $S(t)$ and $S(t')$ that are uncorrelated, e.g. $S(0)$ and $S(T)$, since the dependence lasts for length one.

2.4.1 Glaz approximation for $\mathcal{P}_{\mathbb{X}}(M, h)$

The approximation for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ developed in [30, 34, 137, 138] and discussed in the introduction of this chapter is as follows.

Approximation 5. For $M \geq 2L$ (that is, $T = M/L \geq 2$), the Glaz approximation for the BCP (2.2.4) is

$$\mathcal{P}_{\mathbb{X}}(M, h) \simeq 1 - F_{\mathbb{X}}(2L, h) \left[\frac{F_{\mathbb{X}}(2L, h)}{F_{\mathbb{X}}(L, h)} \right]^{T-2}, \quad (2.4.1)$$

where to approximate the first-passage probabilities $F_{\mathbb{X}}(L, h)$ and $F_{\mathbb{X}}(2L, h)$ defined in (2.2.6), which are $L+1$ and $2L+1$ dimensional integrals respectively, it is advised to use the so-called ‘GenzBretz’ algorithm for numerical evaluation of multivariate normal probabilities; see [28, 29].

Approximation 5 is very accurate. However, its computational cost is also high, especially for large L . Moreover, the main option in the ‘GenzBretz’ package requires the use of Monte-Carlo simulations so that for reliable estimation of high-dimensional integrals one needs to make a lot of averaging; see Section 2.6.1 and 2.8 for more discussion on these issues.

2.4.2 Shepp’s formulas

Define the conditional first-passage probability

$$F(T, h | x) := \Pr \left\{ S(t) < h \text{ for all } t \in [0, T] \mid S(0) = x \right\}. \quad (2.4.2)$$

Since $F(T, h | x) = 0$ for $x > h$, for the unconditional first-passage probability $F(T, h)$ we have $F(T, h) = \int_{-\infty}^h F(T, h | x) \varphi(x) dx$.

The result of [113, p.949] states that if $T = n$ is a positive integer then

$$F(n, h | x) = \frac{1}{\varphi(x)} \int_{D_x} \det[\varphi(y_i - y_{j+1} + h)]_{i,j=0}^n dy_2 \dots dy_{n+1} \quad (2.4.3)$$

where $y_0 = 0, y_1 = h - x, D_x = \{y_2, \dots, y_{n+1} \mid h - x < y_2 < y_3 < \dots < y_{n+1}\}$. For non-integer $T \geq 1$, the exact formula for $F(T, h | x)$ is even more complex (the integral has the dimension $\lceil 2T \rceil$) and completely impractical for computing $P(T, h)$ with $T > 2$, see [113, p.950].

For $n = 1$, we obtain

$$\begin{aligned} F(1, h) &= \int_{-\infty}^h \int_{-x-h}^{\infty} \det \begin{bmatrix} \varphi(x) & \varphi(-x_2 - h) \\ \varphi(h) & \varphi(-x - x_2) \end{bmatrix} dx_2 dx \\ &= \Phi(h)^2 - \varphi(h)[h\Phi(h) + \varphi(h)]. \end{aligned} \quad (2.4.4)$$

For $n = 2$, (2.4.3) yields

$$F(2, h) = \int_{-\infty}^h \int_{-x-a}^{\infty} \int_{x_2-a}^{\infty} \det \begin{bmatrix} \varphi(x) & \varphi(-x_2 - a) & \varphi(-x_3 - 2a) \\ \varphi(a) & \varphi(-x - x_2) & \varphi(-x - a - x_3) \\ \varphi(x_2 + 2a + x) & \varphi(a) & \varphi(x_2 - x_3) \end{bmatrix} dx_3 dx_2 dx. \quad (2.4.5)$$

The three-dimensional integral in (2.4.5) can be reduced to a one-dimensional, see (2.4.20) below with $h_L = h$.

2.4.3 An alternative representation of the Shepp’s formula (2.4.3)

Set $s_i = h + y_i - y_{i+1}$ ($i = 0, 1, \dots, n$) with $s_0 = x, y_0 = 0, y_1 = h - x$. It follows from Shepp’s proof of (2.4.3) that s_0, s_1, \dots, s_n have the meaning of the values of the process $S(t)$ at the times $t = 0, 1, \dots, n$: $S(i) = s_i$ ($i = 0, 1, \dots, n$). The range of the variables s_i is $(-\infty, h)$. Changing the variables in (2.4.3), we obtain

$$F(n, h | x) = \frac{1}{\varphi(x)} \int_{-\infty}^h \dots \int_{-\infty}^h \det[\varphi(s_i + a_{i,j})]_{i,j=0}^n ds_1 \dots ds_n, \quad (2.4.6)$$

where

$$a_{i,j} = y_{i+1} - y_{j+1} = \begin{cases} 0 & \text{for } i = j \\ (i - j)h - s_{j+1} - \dots - s_{i+1} & \text{for } i > j \\ (i - j)h + s_{i+1} + \dots + s_j & \text{for } i < j. \end{cases}$$

2.4.4 Joint density for the values $\{S(i)\}$ and associated transition densities

From (2.4.6), we obtain the following expression for the joint probability density function for the values $S(0), S(1), \dots, S(n)$ under the condition $S(t) < h$ for all $t \in [0, n]$:

$$p(s_0, s_1, \dots, s_n) = \frac{1}{\varphi(s_0)F(n, h | s_0)} \det[\varphi(s_i + a_{i,j})]_{i,j=0}^n. \quad (2.4.7)$$

From this formula, we can derive the transition density from $s_0 = x$ to s_n conditionally $S(t) < h, \forall t \in [0, n]$:

$$q_h^{(0,n)}(x \rightarrow s_n) = \frac{1}{\varphi(x)} \int_{-\infty}^h \dots \int_{-\infty}^h \det[\varphi(s_i + a_{i,j})]_{i,j=0}^n ds_1 \dots ds_{n-1}. \quad (2.4.8)$$

For this transition density, $\int_{-\infty}^h q_h^{(0,n)}(x \rightarrow z) dz = F(n, h | x)$. Moreover, since $S(0) \sim N(0, 1)$, the non-normalized density of $S(n)$ under the condition $S(t) < h$ for all $t \in [0, n]$ is

$$p_h^{(0,n)}(z) := \int_{-\infty}^h q_h^{(0,n)}(x \rightarrow z) \varphi(x) dx \quad (2.4.9)$$

with $z < h$ and $\int_{-\infty}^h p_h^{(0,n)}(z) dz = F(n, h)$. In the case $n = 1$, (2.4.8) gives for $z = s_1 < h$:

$$q_h^{(0,1)}(x \rightarrow z) = \frac{1}{\varphi(x)} \det \begin{pmatrix} \varphi(x) & \varphi(x-h+z) \\ \varphi(h) & \varphi(z) \end{pmatrix} = \varphi(z) \left[1 - e^{-(h-z)(h-x)} \right]. \quad (2.4.10)$$

From this and (2.4.9) we get

$$p_h^{(0,1)}(z) = \int_{-\infty}^h q_h^{(0,1)}(x \rightarrow z) \varphi(x) dx = \Phi(h) \varphi(z) - \Phi(z) \varphi(h)$$

with $z < h$ and $\int_{-\infty}^h p_h^{(0,n)}(z) dz = F(1, h)$.

Rather than just recovering the transition density from $s_0 = x$ to s_n , we can also use (2.4.7) and (2.4.9) to obtain the transition density from $x = s_j$ to $z = s_n$, $0 < j < n$, under the condition $S(t) < h$ for all $t \in [0, n]$:

$$q_h^{(j,n)}(x \rightarrow z) = \frac{1}{p_h^{(0,j)}(z)} \int_{-\infty}^h \dots \int_{-\infty}^h \det[\varphi(s_i + a_{i,j})]_{i,j=0}^n ds_0 ds_1 \dots ds_{j-1} ds_{j+1} \dots ds_{n-1}, \quad (2.4.11)$$

where $s_j = x$ and $s_n = z$. For $j = 1$ and $n = 2$ we obtain the transition density from $x = s_1$ to $z = s_2$ under the condition $S(t) < h$ for all $t \in [0, 2]$:

$$\begin{aligned} q_h^{(1,2)}(x \rightarrow z) &= \frac{1}{p_h^{(0,1)}(z)} \int_{-\infty}^h \det \begin{pmatrix} \varphi(s_0) & \varphi(s_0-h+x) & \varphi(s_0-2h+x+z) \\ \varphi(h) & \varphi(x) & \varphi(x+z-h) \\ \varphi(2h-x) & \varphi(h) & \varphi(z) \end{pmatrix} ds_0 \\ &= \frac{1}{\Phi(h)\varphi(x) - \Phi(x)\varphi(h)} \det \begin{pmatrix} \Phi(h) & \Phi(x) & \Phi(x+z-h) \\ \varphi(h) & \varphi(x) & \varphi(x+z-h) \\ \varphi(2h-x) & \varphi(h) & \varphi(z) \end{pmatrix}. \end{aligned} \quad (2.4.12)$$

2.4.5 Correcting Shepp's formula (2.4.3) for discrete time

2.4.5.1 Rewriting (2.4.3) in terms of the Brownian motion

Let $W(t)$ be the standard Brownian Motion process on $[0, \infty)$ with $W(0) = 0$ and $\mathbb{E}W(t)W(s) = \min(t, s)$. Recall the conditional probability $F(T, h | x)$ defined in (2.4.2). Suppose $T \geq 1$ is an integer and define the event

$$\begin{aligned} \Omega &= \{W(t) < W(t+1) + h < W(t+2) + 2h < \dots < W(t+T) + Th, \forall 0 \leq t \leq 1\} \\ &= \{W(t) - W(t+1) < h, \dots, W(t+T-1) - W(t+T) < h, \forall 0 \leq t \leq 1\}. \end{aligned}$$

If $W(i) = x_i, i = 0, 1, \dots, T+1$, we obtain from [113, p.948]

$$\begin{aligned} F(T, h | x) &= \int \dots \int \Pr\{\Omega \mid W(i) = x_i, i = 0, 1, \dots, T+1, W(0) = 0, W(0) - W(1) = x\} \\ &\quad \times \Pr\{W(i) \in dx_i, i = 0, 1, \dots, T+1 \mid W(0) = 0, W(0) - W(1) = x\}. \end{aligned} \quad (2.4.13)$$

It follows from the proof of (2.4.3) that to correct (2.4.13) for discrete time, one must correct the following probability for discrete time

$$\begin{aligned} &\Pr\{\Omega \mid W(i) = x_i, i = 0, 1, 2, \dots, T+1, W(0) = 0, W(0) - W(1) = x\} \\ &= \Pr\{\sqrt{2}W_1(t) < h, \dots, \sqrt{2}W_T(t) < h, \forall 0 \leq t \leq 1 \mid W(i) = x_i, i = 0, 1, \dots, T+1, \\ &\quad W(0) = 0, W(0) - W(1) = x\} \end{aligned} \quad (2.4.14)$$

where $W_i(t) = \frac{\sqrt{2}}{2}[W(t+i-1) - W(t+i)], i = 1, 2, \dots, T$. Due to the conditioning on the rhs of (2.4.14), the processes $W_i(t)$ can be treated as independent Brownian motion processes. Therefore, the independent increments of the Brownian motion means correcting formula (2.4.3) for discrete time is equivalent to correcting the probability $\Pr(\sqrt{2}W(t) < h, \forall 0 \leq t \leq 1)$ for discrete time.

2.4.5.2 Discretised Brownian motion

To correct the probability $\Pr(\sqrt{2}W(t) < h, \forall 0 \leq t \leq 1)$ for discrete time, we can use the results of Section 2.3.3.2. Using the notation of that section, let $Z = 1$ and $M = L$. This results in $\epsilon = 1/L$ and $t'_n = n\epsilon \in [0, 1], n = 0, 1, \dots, L$. Make the following slight modification of the stopping rule $\tau_{W,h,b}$:

$$\hat{\tau}_{W,h,b} = \inf\{t'_n : \sqrt{2}W(t'_n) \geq h\} = \tau_{W,h/\sqrt{2},b} \quad (2.4.15)$$

and consider the problem of approximating

$$\Pr(\hat{\tau}_{W,h,b} > 1) = \Pr\left(\sqrt{2}W(t'_n) < h \text{ for all } t'_n \in \{0, \epsilon, \dots, L\epsilon = 1\}\right). \quad (2.4.16)$$

With these choice of parameters, as $L \rightarrow \infty$ the piecewise linear continuous-time process $W^\epsilon(t), t \in [0, 1]$, defined in Section 2.3.3.2 converges to $W(t)$ on $[0, 1]$ and so we can refer to $W(t'_n)$ as discretised Brownian motion. We have the following connection between $\sqrt{2}W(t'_n)$ and the random walk Y_n :

$$\sqrt{2}W(t'_n) = \sqrt{2\epsilon}Y_n = \frac{\sqrt{2}}{\sqrt{L}}Y_n, n = 1, 2, \dots, M,$$

where we recall Y_n is a random walk of independent standard normal random variables. Then by using (2.3.11), we approximate the expected excess over the boundary for the process $\sqrt{2}W(t'_n)$ by

$$\omega_L := \frac{0.82}{\sqrt{L}} \simeq \frac{\sqrt{2}\rho}{\sqrt{L}}.$$

Therefore $\omega_L \simeq \sqrt{2}\rho_L$, where ρ_L is defined in (2.3.17). We have deliberately rounded the value $\sqrt{2}\rho \simeq 0.8239\dots$ to 0.82 as for small h and small L it provides marginally better approximation (2.4.18).

Thus, using the same methodology as D. Siegmund and Section 2.3.3.2, to correct the probability $\Pr(\sqrt{2}W(t) < h, \forall 0 \leq t \leq 1)$ for discrete time and obtain an accurate approximation for $\Pr(\hat{\tau}_{W,h,b} > 1)$, we translate the barrier h by the discrete time correction factor ω_L .

2.4.5.3 Corrected version of (2.4.3)

Set $h_L = h + \omega_L$. To correct (2.4.3) for discrete time we substitute the barrier h with h_L . From this and the relation $F(T, h) = \int_{-\infty}^h F(T, h | x)\varphi(x)dx$, the discrete-time corrected form of $F(T, h)$ is

$$\begin{aligned} F(T, h, h_L) &:= \int_{-\infty}^h F(T, h_L | x)\varphi(x)dx \\ &= \int_{-\infty}^h \int_{D_x} \det[\varphi(y_i - y_{j+1} + h_L)]_{i,j=0}^T dy_2 \dots dy_{T+1} dx, \end{aligned} \quad (2.4.17)$$

where $y_0 = 0, y_1 = h_L - x$, and $D_x = \{y_2, \dots, y_{T+1} | h_L - x < y_2 < y_3 < \dots < y_{T+1}\}$.

2.4.5.4 A generic approximation involving corrected Shepp's formula

Approximation 6. For $T = M/L$ an integer with $T \geq 1$, the CDA for the BCP (2.2.4) is

$$\mathcal{P}_{\mathbb{X}}(M, h) \cong P(T, h, h_L) := 1 - F_L(T, h, h_L), \quad (2.4.18)$$

where $F_L(T, h, h_L)$ is given in (2.4.17).

Whilst Approximation 6 is very accurate (see the next subsection), computation of $P(T, h, h_L)$ requires numerical evaluation of a $T + 1$ dimensional integral which is impractical for large T . To overcome this, in Section 2.5.2 we develop approximations that can be easily used for any $T > 0$ (which is not necessarily integer).

2.4.5.5 Particular cases: $T = 1$ and $T = 2$

For $T = 1$, evaluation of (2.4.17) yields

$$F(1, h, h_L) = \Phi(h)\Phi(h_L) - \varphi(h_L)[h\Phi(h) + \varphi(h)]. \quad (2.4.19)$$

This approximation appears simpler than (2.3.17) and is derived under a completely different approach (no use of conditional Gauss Markov properties). For

$T = 1$, our recommendation is to use the simpler Approximation 6 obtained with the use of (2.4.19). For $T < 1$, we recommend the use of (2.3.16).

For $T = 2$, (2.4.17) can be expressed (after some manipulations) as follows:

$$\begin{aligned}
F(2, h, h_L) = & \tag{2.4.20} \\
& \frac{\varphi^2(h_L)}{2} [(h^2 - 1 + \sqrt{\pi}h)\Phi(h) + (h + \sqrt{\pi})\varphi(h)] - \varphi(h_L)\Phi(h_L) [(h + h_L)\Phi(h) + \varphi(h)] \\
& + \Phi(h)\Phi^2(h_L) + \int_0^\infty \Phi(h-y) \left[\varphi(h_L + y)\Phi(h_L - y) - \sqrt{\pi}\varphi^2(h_L)\Phi(\sqrt{2}y) \right] dy.
\end{aligned}$$

Recall from Chapter 1 that in this thesis we treat Φ as if it is explicit and not an integral. This is because $\Phi(x)$ can be easily evaluated by all statistical software. Only a one-dimensional integral has to be numerically evaluated for computing $F(2, h, h_L)$.

2.4.5.6 Simulation study

In this section, we assess the quality of the approximations (2.4.19) and (2.4.20) as well as the sensitivity of the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ to the value of L . In Figures 2.3 and 2.4, the black dashed line corresponds to the empirical values of the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ (for $T = M/L = 1, 2$) computed from 100,000 simulations with different values of L and M (for given L and M , we simulate $L + M$ normal random variables 100,000 times). The solid red line corresponds to Approximation 6. The axis are: the x -axis shows the value of the barrier h in Figure 2.3 and value of L in Figure 2.4; the y -axis denotes the probabilities of reaching the barrier. The graphs, therefore, show the empirical probabilities of reaching the barrier h (for the dashed line) and values of considered approximations for these probabilities. From these graphs we can conclude that Approximation 6 is very accurate, at least for $T = 1, 2$. Furthermore, Approximation 6 is uniformly accurate in h . We can also conclude that the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ is very sensitive to the value of L . From Figure 2.4 we can observe a counter-intuitive fact that even for very high value $L = 1000$, the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ is not even close to $P(T, h)$. This may be explained by the fact that for any fixed T and h , the inaccuracy $|\mathcal{P}_{\mathbb{X}}(M, h) - P(T, h)|$ decreases with the rate const/\sqrt{L} as $L \rightarrow \infty$.

2.4.6 The Glaz-Shepp-Siegmund approximation

Combining (2.4.1) and the approximation (2.4.18) for Shepp's formula (2.4.3), we arrive at the following approximation to which we name the 'Glaz-Shepp-Siegmund approximation'.

Approximation 7. For all $T = M/L > 0$,

$$\mathcal{P}_{\mathbb{X}}(M, h) \simeq 1 - F(2, h, h_L) \cdot \mu_L(h)^{T-2} \quad \text{with } \mu_L(h) = \frac{F(2, h, h_L)}{F(1, h, h_L)}, \tag{2.4.21}$$

where $F(1, h, h_L)$ and $F(2, h, h_L)$ are defined in (2.4.19) and (2.4.20) respectively.

Approximations 5 and 7 look similar but computing Approximation 5 is very hard

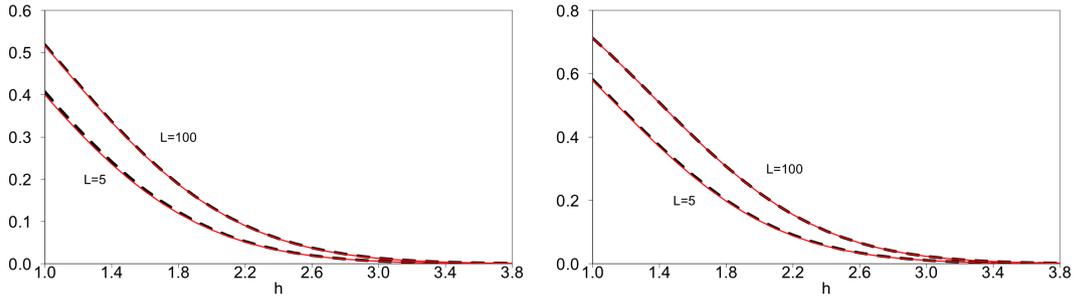


Figure 2.3: Empirical probabilities of reaching the barrier h (dashed black) and corresponding versions of Approximation 6 (solid red). Left: $T = 1$ with (a) $L = M = 5$ and (b) $L = M = 100$. Right: $T = 2$ with (a) $L = 5, M = 10$ and (b) $L = 100, M = 200$.

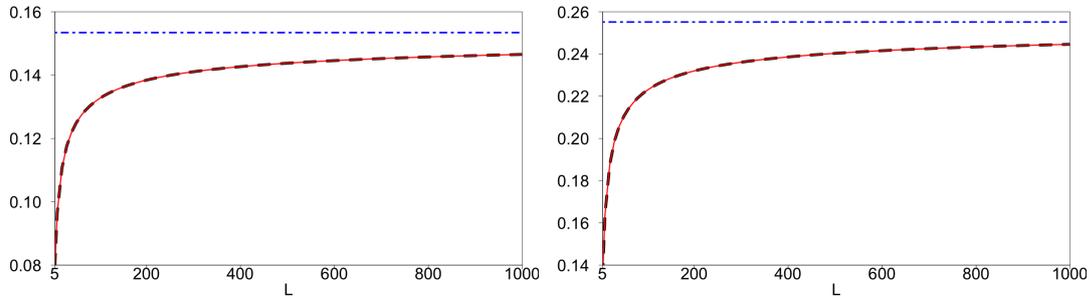


Figure 2.4: Empirical probabilities of reaching the barrier $h = 2$ as a function of L (dashed black), uncorrected diffusion approximation $P(T, 2)$ (dot-dashed blue) and corresponding version of $P_L(T, h)$, which is Approximation 2 (solid red). Left: $M = L$ ($T = 1$). Right: $M = 2L$ ($T = 2$).

and Approximation 7 is very easy (only a one-dimensional integral should be numerically computed).

Note that we have defined Approximation 7 to hold for all $T > 0$. This differs from the original Glaz approximation given in Approximation 5 which is defined for $T \geq 2$. For $T \in [1, 2)$, Approximation 7 remains very accurate. This is demonstrated in Figure 2.5 (left) with $L = 10$ and $T = 1.5$ and confirmed by additional numerical studies not shown here. For $T < 1$, this approximation still remains accurate but is less accurate than Approximation 4, which has been designed to operate specifically for this choice of $T < 1$. This is reflected in Figure 2.5 (right) where we have set $L = 10$ and $T = 0.2$. In these figures, Approximation 7 is depicted with a solid red line, Approximation 4 is depicted with a solid blue line and the dashed black line corresponds to Monte Carlo simulations calculated with 100,000 repetitions.

A more comprehensive assessment of the accuracy of Approximation 7 is provided in Section 2.6. Defining Approximation 7 to hold for all $T > 0$ will be beneficial when approximating $\text{ARL}_H(\mathbb{S}) = \text{ARL}_h(\mathbb{X})$, which will become evident in Section 2.8.

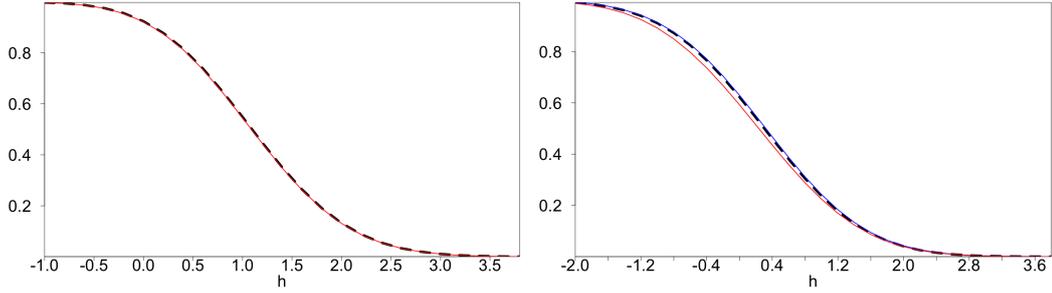


Figure 2.5: Empirical probabilities of reaching the barrier h (dashed black), Approximation 4 (solid blue) and Approximation 7 (solid red). Left: $T = 1.5$ with $L = 10$. Right: $T = 0.2$ with $L = 10$.

2.5 Approximations for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ through eigenvalues of integral operators; $M > L$

The approximations derived in this section still apply to the case where M is larger than L and thus $T = M/L > 1$. The form of approximations in section utilise theory on eigenvalues of integral operators.

2.5.1 Continuous time: approximations for $F(T, h)$

Let m be a positive integer, and $q(x \rightarrow z)$ be the transition density $q_h^{(m-1, m)}(x \rightarrow z)$ defined by (2.4.10) for $m = 1$ (2.4.12) for $m = 2$ and (2.4.11) for $m > 2$.

Let us approximate the distributions of the values $s_i = S(i)$ for integral $i > m$ in the following way. Let $p_i(x)$ be the density of $S(i)$ under the condition that $S(t)$ does not reach h for $t \in [0, i]$. By ignoring the past values of $S(t)$ in $[0, i]$, the non-normalized density of $S(i+1)$ under the conditions that $S(i) \sim p_i(x)$ and $S(t)$ does not reach h for $t \in [i, i+1]$ is

$$\tilde{p}_{i+1}(x) = \int_{-\infty}^h q_h(x \rightarrow z) p_i(y) dy, \text{ for } x < h. \quad (2.5.1)$$

We can then define $p_{i+1}(x) = \tilde{p}_i(x)/c_i$, $x < h$, where $c_i = \int_{-\infty}^h \tilde{p}_i(x) dx$. We then replace formula (2.5.1) with

$$\tilde{p}_i(x) = \int_{-\infty}^h q_h(x \rightarrow z) p(y) dy, \text{ for } x < h, \quad (2.5.2)$$

where $p(x)$ is an eigenfunction of the integral operator with kernel (2.4.10) corresponding to the maximum eigenvalue $\lambda_m(h)$:

$$\lambda_m(h) p(x) = \int_{-\infty}^h p(y) q_h^{(m-1, m)}(x \rightarrow z) dy, \text{ } x < h. \quad (2.5.3)$$

This eigenfunction $p(x)$ is a probability density on $(-\infty, h]$ with $p(x) > 0$ for all $x \in (-\infty, h)$ and $\int_{-\infty}^h p(x) dx = 1$. Moreover, the maximum eigenvalue $\lambda_m(h)$ of

the operator with kernel $K(x, y) = q_h^{(m-1, m)}(x \rightarrow z)$ is simple and positive. The fact that such maximum eigenvalue $\lambda_m(h)$ is simple and real (and hence positive) and the eigenfunction $p(x)$ can be chosen as a probability density follows from the Ruelle-Krasnoselskii-Perron-Frobenius theory of bounded linear positive operators, see e.g. Theorem XIII.43 in [107].

Using (2.5.2) and (2.5.3), we derive recursively: $F(i+1, h) \simeq F(i, h)\lambda_m(h)$ ($i = m, m+1, \dots$). By induction, for any integer $T \geq m$ we then have

$$F(T, h) \simeq F(m, h) \cdot [\lambda_m(h)]^{T-m}. \quad (2.5.4)$$

The approximation (2.5.4) can be used for any $T > 0$ which is not necessarily an integer. The most important particular cases of (2.5.4) are with $m = 1$ and $m = 2$. In these two cases, the kernel $q_h^{(m-1, m)}(x \rightarrow z)$ and hence the approximation (2.5.4) will be corrected for discrete time in the next section.

2.5.2 Correcting approximation (2.5.4) for discrete time

To correct the approximation (2.5.4) for discrete time we need to correct: (a) the first-passage probability $F(m, h)$ and (b) the kernel $q_h^{(m-1, m)}(x \rightarrow z)$. The discrete-time correction of $F(m, h)$ can be done using $F_L(m, h, h_L)$ from (2.4.17) so that what is left is to correct the kernel $q_h^{(m-1, m)}(x \rightarrow z)$ and hence $\lambda_m(h)$.

2.5.2.1 Correcting the transition kernels for discrete time

As explained in Section 2.4.5, to make a discrete-time correction in the Shepp's formula (2.4.3) we need to replace the barrier h with $h_L = h + \omega_L$ in all places except for the upper bound for the initial value $S(0)$. Therefore, using the notation of Section 2.4.3, the joint probability density function for the values $S(0), S(1), \dots, S(m)$ under the condition $S(t) < h$ for all $t \in [0, m]$ corrected for discrete time is:

$$\hat{p}(s_0, s_1, \dots, s_m) = \frac{1}{\varphi(s_0)F(m, h | s_0)} \det[\varphi(s_i + \hat{a}_{i,j})]_{i,j=0}^m \quad (2.5.5)$$

with $-\infty < s_0 < h$, $-\infty < s_j < h_L$ ($j = 1, \dots, m$),

$$\hat{a}_{i,j} = y_{i+1} - y_{j+1} = \begin{cases} 0 & \text{for } i = j \\ (i-j)h_L - s_{j+1} - \dots - s_{i+1} & \text{for } i > j \\ (i-j)h_L + s_{i+1} + \dots + s_j & \text{for } i < j. \end{cases}$$

This provides the discrete-time corrected transition density from $s_0 = x$ to s_m conditionally $S(t) < h$, $\forall t \in [0, m]$:

$$q_{h_L}^{(0, m)}(x \rightarrow s_m) = \frac{1}{\varphi(x)} \int_{-\infty}^{h_L} \dots \int_{-\infty}^{h_L} \det[\varphi(s_i + \hat{a}_{i,j})]_{i,j=0}^m ds_1 \dots ds_{m-1}; \quad (2.5.6)$$

which is exactly (2.4.8) with h_L is substituted for h . In a particular case $m = 1$, the corrected transition density is

$$\begin{aligned} q_{h_L}^{(0, 1)}(x \rightarrow s_1) &= \frac{1}{\varphi(x)} \det \begin{pmatrix} \varphi(x) & \varphi(x - h_L + s_1) \\ \varphi(h_L) & \varphi(s_1) \end{pmatrix} \\ &= \varphi(s_1) \left[1 - e^{-(h_L - s_1)(h_L - x)} \right] \end{aligned} \quad (2.5.7)$$

with $s_1 < h_L$.

Let us now make the discrete-time correction of the transition density $q_h^{(1,2)}(x \rightarrow z)$. Denote by $p_{h,L}^{(0,1)}(z)$, $z < h$, the non-normalized density of $S(1)$ under the condition $S(t) < h$ for all $t \in [0, 1]$ corrected for discrete time; it satisfies $\int_{-\infty}^h p_{h,L}^{(0,1)}(z) dz = F(1, h, h_L)$. Using (2.5.7), we obtain

$$p_{h,L}^{(0,1)}(z) = \int_{-\infty}^h q_{h_L}^{(0,1)}(x \rightarrow z) \varphi(x) dx = \varphi(z) \Phi(h) - \varphi(h_L) \Phi(h - h_L + z).$$

From (2.5.5) and (2.5.7), the transition density from $x = s_1$ to $z = s_2$ under the condition $S(t) < h$ for all $t \in [0, 2]$ corrected for discrete time (the corrected form of (2.4.12)) is given by

$$\begin{aligned} q_{h,L}^{(1,2)}(x \rightarrow z) &= \frac{1}{p_{h,L}^{(0,1)}(x)} \int_{-\infty}^h \det \begin{pmatrix} \varphi(s_0) & \varphi(s_0 - h_L + x) & \varphi(s_0 - 2h_L + x + z) \\ \varphi(h_L) & \varphi(x) & \varphi(x + z - h_L) \\ \varphi(2h_L - x) & \varphi(h_L) & \varphi(z) \end{pmatrix} ds_0 \\ &= \frac{1}{p_{h,L}^{(0,1)}(x)} \det \begin{pmatrix} \Phi(h) & \Phi(h - h_L + x) & \Phi(h - 2h_L + x + z) \\ \varphi(h_L) & \varphi(x) & \varphi(x + z - h_L) \\ \varphi(2h_L - x) & \varphi(h_L) & \varphi(z) \end{pmatrix} \end{aligned} \quad (2.5.8)$$

Unlike the transition density (2.5.6) (and (2.5.7) in the particular case $m = 1$), which only depends on h_L and not on h , the transition density $q_{h,L}^{(1,2)}(x \rightarrow z)$ depends on both h and h_L and hence the notation. The dependence on h has appeared from integration over the $s_0 \in (-\infty, h)$.

2.5.2.2 Approximations for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$

With discrete-time corrected transition densities $q_h^{(0,1)}(x \rightarrow z)$ and $q_h^{(1,2)}(x \rightarrow z)$, we obtain the corrected versions of the approximations (2.5.4).

Approximation 8: $\mathcal{P}_{\mathbb{X}}(M, h) \simeq 1 - F(1, h, h_L) \cdot [\lambda_{L,1}(h)]^{T-1}$, where $T = M/L$, $F(1, h, h_L)$ is given in (2.4.19) and $\lambda_{L,1}(h)$ is the maximal eigenvalue of the integral operator with kernel $K(x, z) = q_{h_L}^{(0,1)}(x \rightarrow z)$ defined in (2.5.7).

Approximation 9: $\mathcal{P}_{\mathbb{X}}(M, h) \simeq 1 - F(2, h, h_L) \cdot [\lambda_{L,2}(h)]^{T-2}$, where $T = M/L$, $F(2, h, h_L)$ is given in (2.4.20) and $\lambda_{L,2}(h)$ is the maximal eigenvalue of the integral operator with kernel $K(x, z) = q_{h,L}^{(1,2)}(x \rightarrow z)$ defined in (2.5.8).

Similarly to $\lambda_m(h)$ from (2.5.3), the maximum eigenvalues $\lambda_{L,1}(h)$ and $\lambda_{L,2}(h)$ of the operators with kernels $K(x, z) = q_{h_L}^{(0,1)}(x \rightarrow z)$ and $K(x, z) = q_{h,L}^{(1,2)}(x \rightarrow z)$ are simple and positive; the corresponding eigenfunctions $p(x)$ can be chosen as probability densities. Both approximations can be used for any $T > 0$.

In numerical examples below we approximate the eigenvalues $\lambda_{L,k}(h)$ ($k = 1, 2$) using the methodology described in [72], p.154. This methodology is based on the Gauss-Legendre discretization of the interval $[-c, h]$, with some large $c > 0$, into an N -point set x_1, \dots, x_N (the x_i 's are the roots of the N -th Legendre polynomial on $[-c, h]$), and the use of the Gauss-Legendre weights w_i associated with points

x_i ; $\lambda_{L,k}(h)$ and $p(x)$ are then approximated by the largest eigenvalue and associated eigenvector of the matrix $D^{1/2}AD^{1/2}$, where $D = \text{diag}(w_i)$ and $A_{i,j} = K(x_i, x_j)$ with the respective kernel $K(x, z)$. If N is large enough then the resulting approximation to $\lambda_{L,k}(h)$ is arbitrarily accurate. With modern software, computing Approximations 8 and 9 (as well as Approximation 7) with high accuracy takes only milliseconds on a regular laptop.

As discussed in the next section, Approximation 9 is more accurate than Approximation 8, especially for small h ; the accuracies of Approximations 7 and 9 are very similar. Note also that a version of Approximation 8 has been developed in the authors previous work in [82] (not shown in this thesis); this version was based on a different discrete-time approximation (discussed in Section 2.4.5.5) of the continuous-time BCP probability $P(T, h)$.

2.6 Simulation study

2.6.1 Accuracy of approximations for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$

In this section we study the quality of Approximations 8 and 9 for the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ defined in (2.2.4). Approximation 7 is visually indistinguishable from Approximation 9 and is therefore not plotted (see Table 2.2). Without loss of generality, ε_j in (2.1.1) are normal r.v.'s with mean 0 and variance 1. The style of Figure 2.6 is exactly the same as of Figure 2.3 and is described in the beginning of Section 2.4.5.5. In Figure 2.6, the dashed green line corresponds to Approximation 8 and the solid red line corresponds to Approximation 9. Like in Figure 2.3, for Monte Carlo simulations we have take 100,000 repetitions.

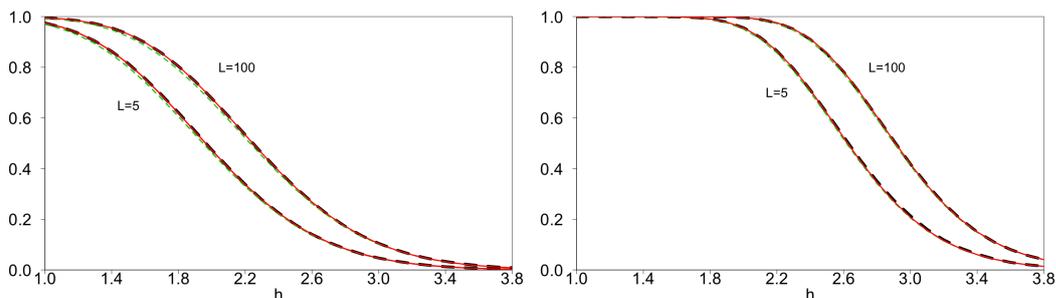


Figure 2.6: Empirical probabilities of reaching the barrier h (dashed black), Approximation 8 (dashed green) and Approximation 9 (solid red). Left: $T = 10$ with (a) $L = 5$ and (b) $L = 100$. Right: $T = 50$ with (a) $L = 5$ and (b) $L = 100$.

From Figure 2.6 we see that the performance of Approximations 8 and 9 is very strong across all values of h even for small L . For small h , Approximation 9 is slightly more accurate than Approximation 8 in view of its better accommodation to the non-Markovian nature of the process $S(t)$.

In Table 2.2, we display the values of $\lambda_{L,1}(h)$, $\lambda_{L,2}(h)$ and $\mu_L(h)$ with $L = 20$ for a number of different h . From this table, we see only a small difference be-

	$h=0$	$h=0.5$	$h=1$	$h=1.5$	$h=2$	$h=2.5$	$h=3$	$h=3.5$	$h=4$
$\lambda_{L,1}(h)$	0.28494	0.46443	0.65331	0.81186	0.91687	0.97090	0.99209	0.99835	0.99974
$\lambda_{L,2}(h)$	0.25744	0.43811	0.63472	0.80239	0.91348	0.97005	0.99195	0.99833	0.99974
$\mu_L(h)$	0.25527	0.43677	0.63432	0.80241	0.91353	0.97007	0.99195	0.99833	0.99974

Table 2.2: Values of $\lambda_{L,1}(h)$, $\lambda_{L,2}(h)$ and $\mu_L(h)$ with $L = 20$ for different h .

tween $\lambda_{L,2}(h)$ and $\mu_L(h)$; this difference is too small to visually differentiate between Approximations 7 and 9 in Figure 2.6.

In Tables 2.3, 2.4 and 2.5 we numerically compare the performance of Approximations 5 and 7 for approximating $\mathcal{P}_{\mathbb{X}}(M, h)$ across different values of L and h . Since Approximation 5 relies on Monte-Carlo methods, we present the average over 100 evaluations and denote this by \bar{x} . We have also provided values for the standard deviation and maximum and minimum of the 100 runs to illustrate the randomised nature of this approximation. These are denoted by s , $\max(x_i)$ and $\min(x_i)$ respectively. The values of $\mathcal{P}_{\mathbb{X}}(M, h)$ presented in the tables below are the empirical probabilities of reaching the barrier h obtained by 10^6 simulations. Approximation 9 is not included in these tables as results are identical to Approximation 7 up to four decimal places.

	$h=2.5$	$h=2.75$	$h=3$	$h=3.25$	$h=3.5$	$h=3.75$	$h=4$
\bar{x}	0.855957	0.627299	0.376337	0.191122	0.086253	0.033769	0.013156
s	0.004127	0.008588	0.013805	0.015181	0.012826	0.008510	0.005131
$\max(x_i) - \bar{x}$	0.010665	0.023748	0.029819	0.027066	0.025629	0.016208	0.011609
$\bar{x} - \min(x_i)$	0.012176	0.021268	0.033211	0.041322	0.041350	0.022650	0.018146
Approx 7	0.854844	0.625113	0.373863	0.188933	0.083981	0.033833	0.012551
$\mathcal{P}_{\mathbb{X}}(M, h)$	0.855429	0.627463	0.376681	0.191625	0.085697	0.034675	0.013116

Table 2.3: Average values from 100 evaluations of Approximation 5 for different h along with maximum and minimum with $L = 5$ and $T = 100$.

	$h=2.5$	$h=2.75$	$h=3$	$h=3.25$	$h=3.5$	$h=3.75$	$h=4$
\bar{x}	0.952007	0.802073	0.554613	0.315085	0.155331	0.066113	0.025608
s	0.001479	0.004856	0.012540	0.015050	0.015160	0.011647	0.008129
$\max(x_i) - \bar{x}$	0.004746	0.013360	0.027078	0.030940	0.033991	0.024111	0.030014
$\bar{x} - \min(x_i)$	0.003662	0.010894	0.031463	0.037715	0.041021	0.043283	0.016997
Approx 7	0.952475	0.802100	0.555109	0.316076	0.153803	0.066438	0.026143
$\mathcal{P}_{\mathbb{X}}(M, h)$	0.952818	0.803078	0.555530	0.315784	0.153446	0.066642	0.026244

Table 2.4: Average values from 100 evaluations of Approximation 5 for different h along with maximum and minimum with $L = 20$ and $T = 100$.

	$h=2.5$	$h=2.75$	$h=3$	$h=3.25$	$h=3.5$	$h=3.75$	$h=4$
\bar{x}	0.979027	0.878031	0.661247	0.402887	0.211894	0.093329	0.039110
s	0.000884	0.005502	0.014418	0.021283	0.018493	0.020459	0.015536
$\max(x_i) - \bar{x}$	0.001995	0.009243	0.039695	0.040615	0.063578	0.064306	0.037958
$\bar{x} - \min(x_i)$	0.002414	0.020613	0.025530	0.093876	0.038484	0.05694	0.033748
Approx. 7	0.979119	0.878481	0.660662	0.405674	0.209313	0.094517	0.038529

Table 2.5: Average values from 100 evaluations of Approximation 5 for different h along with maximum and minimum with $L = 100$ and $T = 100$.

From Tables 2.3, 2.4 and 2.5 we see that with this choice of $T = 100$, the errors of approximating $F_{\mathbb{X}}(2L, h)$ and $F_{\mathbb{X}}(L, h)$ via the 'GenzBretz' algorithm can accumulate and lead to a fairly significant variation of Approximation 5. This demonstrates the need to average the outcomes of Approximation 5 over a significant number of runs, should one desire an accurate approximation. This may require rather high computational cost and run time, especially if L is large. On the other hand, evaluation of Approximation 7 is practically instantaneous for all L . Even for a very small choice of $L = 5$, Table 2.3 shows that Approximation 7 still remains very accurate. As L increases from 5 to 20, Table 2.4 shows that the accuracy of Approximation 7 increases. The averaged Approximation 5 is also very accurate but a larger L appears to produce a larger range for $\max(x_i)$ and $\min(x_i)$ when h is large; this is seen in Table 2.5. Note the empirical values of $\mathcal{P}_{\mathbb{X}}(M, h)$ are not included in Table 2.5 due to the large computational cost.

2.6.2 Approximation for the BCP in the case of non-normal moving sums

Approximations 7, 8 and 9 remain very accurate when then the original ε_i in (2.1.1) are not exactly normal. We consider two cases: (a) ε_i are uniform r.v.'s on $[0,1]$ and (b) ε_i are Laplace r.v.'s with mean zero and scale parameter 1. Simulation results are shown in Figure 2.7; this figure has the same style as figures in Sections 2.4.5.6 and 2.6.1 and Monte Carlo simulations have been performed with 100,000 simulations.

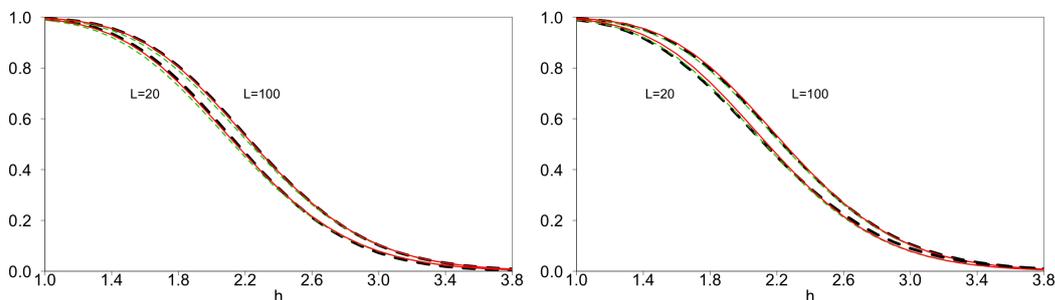


Figure 2.7: Empirical probabilities of reaching the barrier h (dashed black), Approximation 8 (dashed green) and Approximation 9 (solid red). Left: $\varepsilon_i \sim \text{Uniform}[0, 1]$ and $T = 10$ with (a) $L = 20$ and (b) $L = 100$. Right: $\varepsilon_i \sim \text{Laplace}[0, 1]$ and $T = 10$ with (a) $L = 20$ and (b) $L = 100$.

Some selected values used for plots in Figure 2.7 are:

$h = 2, L = 20$: Emp: 0.6045 ± 0.0030 (0.6123 ± 0.0030) [0.5894 ± 0.003]; Ap. 4(5): $0.5921(0.6054)$;
 $h = 2, L = 100$: Emp: 0.6771 ± 0.0029 (0.6801 ± 0.0029) [0.6722 ± 0.003]; Ap. 4(5): $0.6633(0.6775)$;
 $h = 3, L = 20$: Emp: 0.0788 ± 0.0017 (0.0710 ± 0.0016) [0.0915 ± 0.002]; Ap. 4(5): $0.0777(0.0789)$;
 $h = 3, L = 100$: Emp: 0.1039 ± 0.0019 (0.1033 ± 0.0019) [0.1048 ± 0.002]; Ap. 4(5): $0.1022(0.1034)$.

Here we provide means and 95% confidence intervals for the empirical (Emp) values of the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ (with $T = M/L = 10$) computed from 100,000 Monte-Carlo runs of the sequences of the moving sums (2.1.1) with normal (no brack-

ets), uniform (regular brackets) and Laplace (square brackets) distributions for ε_i in (2.1.1). Values of Approximations (Ap.) 8 and 9 are also given.

From Figure 2.7 and associated numbers we can make the following conclusions: (a) the BCP $\mathcal{P}_{\mathbb{X}}(M, h)$ for the case where ε_i in (2.1.1) are uniform is closer to the case where ε_i are normal, than for the case where ε_i have Laplace distribution; (b) as L increases, the probabilities $\mathcal{P}_{\mathbb{X}}(M, h)$ in the cases of uniform and Laplace distributions of ε_i become closer to the BCP for the case of normal ε_i and hence the approximations to the BCP become more precise; (c) accuracy of Approximation 9 is excellent for the case of normal ε_i and remains very good in the case of uniform ε_i ; it is also rather good in the case when ε_i have Laplace distribution; (d) Approximation 8 is slightly less accurate than Approximation 9 (and Approximation 7) for the case of normal and uniform ε_i (this is in a full agreement with discussions in Sections 2.5.2.2 and 2.6.1); however, Approximation 8 is very simple and can still be considered as rather accurate.

2.6.3 Approximation for the BCP in the case of moving weighted sums

The author has also investigated the performance of Approximation 9 (and 7) after introducing particular weights into (2.1.1). The following two ways of incorporating weights has been explored:

- (i) L random weights w_1, w_2, \dots, w_L , with w_i i.i.d. uniform on $[0, 2]$, are associated with a position in the moving window; this results in the moving weighted sum

$$S_{n,w,L} := \sum_{j=n+1}^{n+L} w_{j-n} \varepsilon_j \quad (n = 0, 1, \dots, M);$$

- (ii) $M + L$ random weights w_1, \dots, w_{M+L} are associated with r.v. $\varepsilon_1, \dots, \varepsilon_{M+L}$; here w_j are i.i.d. uniform r.v.'s on $[0, 2]$; this gives the moving weighted sum

$$S_{n,w,L} := \sum_{j=n+1}^{n+L} w_j \varepsilon_j \quad (n = 0, 1, \dots, M).$$

Simulations results are shown in Figure 2.8. In both cases, we have repeated simulations 1,000 times and plotted all the curves representing the BCP as functions of h in grey colour and Approximation 9 for the BCP for the non-weighted case (when all weights $w_j = 1$) as red dashed line. We can see that for both scenarios the Approximation 9 for the BCP in the non-weighted case gives fairly accurate approximation for the weighted BCP. Similar results have been observed for other values of L and T .

2.7 Approximation using Extreme value theory; $M > L$

For M large and significantly greater than L , it is possible to use results from Extreme value theory. For $L = [c \log M]$ with $c > 0$, from Theorem 1.5 in [48] it was proved

$$\lim_{M \rightarrow \infty} \mathcal{P}_{\mathbb{X}}(M, h) = 1 - \exp(-e^{-(h-a_M)/b_M}), \quad (2.7.1)$$

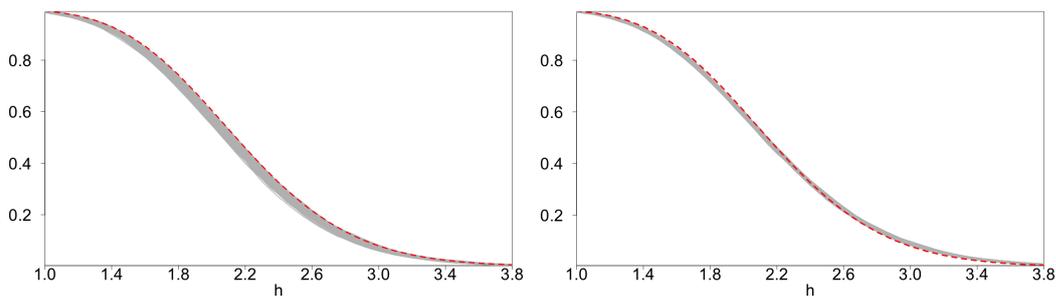


Figure 2.8: BCP for the weighted sums (grey) against Approximation 9 for the BCP for non-weighted moving sums (red dotted line). Left: case (i) with $L = 20, M = 200, T = 10$. Right: case (ii) with $L = 20, M = 200, T = 10$.

where the constants a_M and b_M are given by

$$a_M = \sqrt{2 \log M} + \frac{-1/2 \log \log M + \log((4/c)F(4/c)) - \log(2\sqrt{\pi})}{\sqrt{2 \log M}}, \quad b_M = \frac{1}{\sqrt{2 \log M}}.$$

The function $F(a)$ satisfies

$$F(a) = \frac{p^2(a)}{a}, \quad \text{where } p(a) = \exp\left(-\sum_{k=1}^{\infty} \frac{1}{k} \Pr(Z_k > 0)\right)$$

and $Z_n = \sum_{i=1}^n X_i$ with X_i i.i.d. Gaussian with mean $-a/2$ and variance a . The constant $p(a)$ requires numerical evaluation. We remark that similar findings are in [120]. As a result, for a finite M we make the following approximation.

Approximation 10: *Extreme value approximation for the BCP (2.2.4):* For $M/L > 1$ and $c = L/\log(M)$: $\mathcal{P}_{\mathbb{X}}(M, h) \simeq 1 - \exp(-e^{-(h-a_M)/b_M})$.

Of course, this approximation will only perform well for large M and hence large h , as is shown in Figure 2.9. In this figure, the dashed black line corresponds to $\mathcal{P}_{\mathbb{X}}(M, h)$ obtained by Monte Carlo simulations with 100,000 iterations. The solid red line depicts Approximation 10. The accuracy of this approximation seems comparable to Approximation 1 and Approximation 2 and is poor for small T (see Figure 2.9 right) and clearly not uniform in h as is desired.

2.8 Approximating Average Run Length (ARL)

As discussed in (2.2.10), the diffusion approximation for $\text{ARL}_h(\mathbb{X})/L$ is

$$\text{ARL}_h(S(t)) = \mathbb{E}(\tau_h(S(t))) = \int_0^{\infty} s q(s, h, S(t)) ds. \quad (2.8.1)$$

The diffusion approximation (2.8.1) should be corrected for discrete time; otherwise it is poor, especially for small L . As shown in Section 2.6 and recalling the discussion towards the end of Section 2.4.6, Approximations 7 and 9 are very accurate approximations for $\mathcal{P}_{\mathbb{X}}(M, h)$ for all $T > 0$ ($M > 0$). In the results below, one

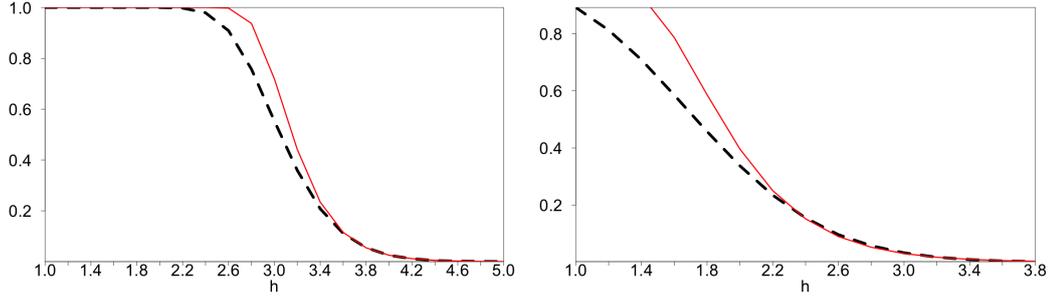


Figure 2.9: Empirical probabilities of reaching the barrier h (dashed black) and Approximation 10 (solid red). Left: $T = 100$ with $L = 20$. Right: $T = 5$ with $L = 10$.

could implement the approximations derived in Section 2.3 for when $T \leq 1$, however the additional complexity is undesirable. We shall use Approximation 7 to formulate our approximations but note that the use of Approximation 9 would provide very similar results.

We define the approximation $\hat{q}(s, h)$ for the probability density function of $\tau_h(\mathbb{X})/L$ by

$$\hat{q}(s, h) = \frac{d}{ds} \{1 - F(2, h, h_L) \cdot \mu_L(h)^{s-2}\} = -F(2, h, h_L) \log(\mu_L(h)) \cdot \mu_L(h)^{s-2},$$

for $s > 0$. The corresponding approximation for $\text{ARL}_h(\mathbb{X})$ is

$$\text{ARL}_h(\mathbb{X}) = \mathbb{E} \tau_h(\mathbb{X}) \cong L \int_0^\infty s \hat{q}(s, h) ds = -\frac{L \cdot F(2, h, h_L)}{\mu_L(h)^2 \log(\mu_L(h))}. \quad (2.8.2)$$

The standard deviation of $\tau_h(\mathbb{X})$, denoted $SD(\tau_h(\mathbb{X}))$, is approximated by:

$$SD(\tau_h(\mathbb{X})) \cong L \left[\int_0^\infty s^2 \hat{q}(s, h) ds - \left(\int_0^\infty s \hat{q}(s, h) ds \right)^2 \right]^{1/2}. \quad (2.8.3)$$

In this chapter, we define ARL in terms of the number of random variables $\xi_{n,L}$ rather than number of random variables ε_j . This means we have to modify the approximation for ARL of [34] by subtracting L . The standard deviation approximation in [34] is not altered.

The Glaz approximations for $\text{ARL}_h(\mathbb{X})$ and $SD(\tau_h(\mathbb{X}))$ are as follows:

$$\mathbb{E}_G(\tau_h(\mathbb{X})) = \sum_{j=L}^{2L} (F_{\mathbb{X}}(j-L, h)) + \frac{F_{\mathbb{X}}(L, h)}{F_{\mathbb{X}}(L, h) - F_{\mathbb{X}}(2L, h)} \sum_{j=1}^L (F_{\mathbb{X}}(L+j, h)), \quad (2.8.4)$$

$$SD_G(\tau_h(\mathbb{X})) = \left[L(L-1) + 2 \sum_{j=L}^{3L} j (F_{\mathbb{X}}(j-L, h)) + \frac{2Lx(3-2x)}{(1-x)^2} \sum_{j=1}^L (F_{\mathbb{X}}(j+L, h)) \right. \\ \left. + \frac{2x}{1-x} \sum_{j=1}^L j (F_{\mathbb{X}}(j+L, h)) + \mathbb{E}_G(\tau_h(\mathbb{X})) - \mathbb{E}_G(\tau_h(\mathbb{X}))^2 \right]^{1/2}, \quad (2.8.5)$$

where $x = F_{\mathbb{X}}(2L, h)/F_{\mathbb{X}}(L, h)$.

In Tables 2.6 and 2.7 we assess the accuracy of the approximations (2.8.2) and (2.8.3) and also Glaz approximations (2.8.4) and (2.8.5). In these tables, the values of $\text{ARL}_h(\mathbb{X})$ and $SD(\tau_h(\mathbb{X}))$ have been calculated using 100,000 simulations. Since the Glaz approximations rely on Monte Carlo methods, in the tables we have reported value $2s$ -confidence intervals computed from 150 evaluations.

h	2	2.25	2.5	2.75	3	3.25	3.5
(2.8.2)	126	217	395	759	1551	3375	7837
(2.8.4)	126 \pm 1	218 \pm 2	394 \pm 5	756 \pm 17	1545 \pm 65	3388 \pm 300	7791 \pm 1100
$\text{ARL}_h(\mathbb{X})$	127	218	396	757	1550	3344	7721

h	2	2.25	2.5	2.75	3	3.25	3.5
(2.8.3)	129	220	397	761	1553	3377	7839
(2.8.5)	129 \pm 1	220 \pm 2	397 \pm 5	758 \pm 17	1549 \pm 65	3389 \pm 300	7793 \pm 1100
$SD(\tau_h(\mathbb{X}))$	129	221	395	758	1550	3341	7716

Table 2.6: Approximations for $\text{ARL}_h(\mathbb{X})$ and $SD(\tau_h(\mathbb{X}))$ with $L = 10$.

h	2	2.25	2.5	2.75	3	3.25	3.5
(2.8.2)	471	791	1392	2587	5099	10695	23918
(2.8.4)	471 \pm 3	791 \pm 7	1393 \pm 25	2597 \pm 75	5101 \pm 270	10708 \pm 1250	24639 \pm 5800
$\text{ARL}_h(\mathbb{X})$	472	792	1397	2588	5085	10749	24131

h	2	2.25	2.5	2.75	3	3.25	3.5
(2.8.3)	485	804	1404	2598	5109	10704	23924
(2.8.5)	481 \pm 3	802 \pm 7	1404 \pm 25	2608 \pm 75	5147 \pm 270	10716 \pm 1250	24649 \pm 5800
$SD(\tau_h(\mathbb{X}))$	485	804	1407	2600	5093	10762	24105

Table 2.7: Approximations for $\text{ARL}_h(\mathbb{X})$ and $SD(\tau_h(\mathbb{X}))$ with $L = 50$.

Tables 2.6 and 2.7 show that the approximations developed in this chapter perform strongly and are similar, for small or moderate h , to the Glaz approximations. For $h \geq 3$, the Glaz approximation produces rather large uncertainty intervals and the uncertainty quickly deteriorates with the increase of h . This is due to the fairly large uncertainty intervals formed by Approximation 5 when approximating $\mathcal{P}_{\mathbb{X}}(M, h)$ with large h and hence small $\mathcal{P}_{\mathbb{X}}(M, h)$, as discussed in Section 2.6.1. The approximations developed in this chapter are deterministic and are much simpler in comparison to the Glaz approximations. Moreover, they do not deteriorate for large h .

2.9 Appendix A: Proof of Lemma 2.2.1

As correlation is invariant under linear transformations, $\text{Corr}(S_{0,L}, S_{k,L}) = \text{Corr}(\xi_0, \xi_k)$. From the definition (2.1.1) we have $\text{Corr}(S_{0,L}, S_{k,L}) = \text{Corr}(S_{n,L}, S_{n+k,L})$. The sum $S_{k,L}$ can be represented as

$$S_{k,L} = S_{0,L} - \sum_{j=1}^k \varepsilon_j + \sum_{j=L+1}^{L+k} \varepsilon_j.$$

Using this representation, we obtain

$$\text{Cov}(S_{0,L}, S_{k,L}) = \underbrace{(\sigma^2 L + \mu^2 L^2)}_{\mathbb{E} S_{0,L}^2} - k\sigma^2 - \underbrace{\mu^2 L^2}_{(\mathbb{E} S_{0,L})^2} = \sigma^2 L - k\sigma^2.$$

Dividing this by $\text{var}(S_{0,L})$, from (2.2.1), we obtain $\text{Corr}(S_{0,L}, S_{k,L}) = 1 - k/L$ in the case $k \leq L$. The case $k > L$ is obvious.

2.10 Appendix B: Derivation of Durbin approximation

We shall initially show $R'(0+) = -1 \neq 0$. We have

$$\left. \frac{\partial R(t, s)}{\partial s} \right|_{s=t+} = R(0+).$$

Using (2.2.7) and the fact that $\Delta = 1/L$, we have

$$R'(0+) = \lim_{L \rightarrow \infty} \frac{R(\Delta) - R(0)}{\Delta} = - \lim_{L \rightarrow \infty} \frac{L}{L} = -1.$$

The Durbin approximation for $q(t, h, \zeta_t)$ can be written as

$$q(t, h, S(t)) \cong b_0(t, h) f(t, h),$$

where

$$f(t, h) = \frac{1}{\sqrt{2\pi R(t, t)}} e^{-\frac{h^2(t)}{2R(t, t)}}, \quad b_0(t, h) = - \frac{h(t)}{R(t, t)} \left. \frac{\partial R(s, t)}{\partial s} \right|_{s=t+} - \frac{dh(t)}{dt}.$$

In view of (2.2.9) the related approximation for the first passage probability $P(T, h)$ is

$$P(T, h) \cong \int_0^T b_0(t, h) f(t, h) dt.$$

In the case when the threshold $h(t) = h$ is constant, using Lemma 2.2.2 we obtain

$$b_0(t, h) = -hR'(0+) = h, \quad q(t, h, S(t)) \cong \frac{h}{\sqrt{2\pi}} e^{-h^2/2}$$

and therefore we obtain the following approximation.

$$\mathcal{P}_{\mathbb{X}}(M, h) \cong P(T, h) \cong \frac{hT}{\sqrt{2\pi}} e^{-h^2/2}.$$

2.11 Appendix C: Derivation of (2.3.17)

As $M = L$, $P_\rho(L, h) = \int_{-\infty}^h (1 - \Phi(b + \hat{a}) + e^{-2\hat{a}b} \Phi(b - \hat{a})) \varphi(x_0) dx_0 + 1 - \Phi(h)$.

Using the fact $\hat{a} = (h - x_0)/2 + \rho_L$ and $b = (h + x_0)/2$, we obtain:

$$\begin{aligned} P_\rho(L, h) &= 1 - \int_{-\infty}^h \Phi(h + \rho_L) \varphi(x_0) dx_0 + \int_{-\infty}^h e^{-h^2/2 + x_0^2/2 - \rho_L h - \rho_L x_0} \Phi(x_0 - \rho_L) \varphi(x_0) dx_0 \\ &= 1 - \Phi(h + \rho_L) \Phi(h) + \varphi(h) e^{-\rho_L h} \int_{-\infty}^h e^{-\rho_L x_0} \int_{-\infty}^{x_0 - \rho_L} \varphi(z) dz dx_0. \end{aligned}$$

Making the substitution $k = z + \rho_L$ in the rightmost integral, we obtain

$$\varphi(h)e^{-\rho_L h} \int_{-\infty}^h \int_{-\infty}^{x_0} e^{-\rho_L x_0} \varphi(k - \rho_L) dk dx_0.$$

By then changing the order of integration:

$$\varphi(h)e^{-\rho_L h} \int_{-\infty}^h \int_k^h e^{-\rho_L x_0} \varphi(k - \rho_L) dx_0 dk = \frac{\varphi(h)e^{-\rho_L h}}{\rho_L} \int_{-\infty}^h (e^{-\rho_L k} - e^{-\rho_L h}) \varphi(k - \rho_L) dk.$$

By expanding the brackets, we obtain:

$$\begin{aligned} \frac{\varphi(h)e^{-\rho_L h}}{\rho_L} \int_{-\infty}^h (e^{-\rho_L k} - e^{-\rho_L h}) \varphi(k - \rho_L) dk &= \frac{\varphi(h)e^{-\rho_L h - \rho_L^2/2}}{\rho_L} \int_{\infty}^h \varphi(k) dk \\ &- \frac{\varphi(h)e^{-2\rho_L h}}{\rho_L} \int_{\infty}^h \varphi(k - \rho_L) dk \\ &= \frac{\varphi(h + \rho_L)}{\rho_L} \Phi(h) - \frac{\varphi(h)e^{-2\rho_L h}}{\rho_L} \Phi(h - \rho_L). \end{aligned}$$

Thus we obtain the required:

$$P_\rho(L, h) = 1 - \Phi(h + \rho_L)\Phi(h) + \frac{\varphi(h + \rho_L)}{\rho_L} \Phi(h) - \frac{\varphi(h)e^{-2h\rho_L}}{\rho_L} \Phi(h - \rho_L).$$

Chapter 3

Approximating Shepp's constants for the Slepian process

Abstract

Slepian process $S(t)$ is a stationary Gaussian process with zero mean and covariance $\mathbb{E}S(t)S(t') = \max\{0, 1 - |t - t'|\}$. For any $T \geq 0$ and real h , define $F_T(h) = \Pr\{\max_{t \in [0, T]} S(t) < h\}$ and the constants $\Lambda(h) = -\lim_{T \rightarrow \infty} \frac{1}{T} \log F_T(h)$ and $\lambda(h) = \exp\{-\Lambda(h)\}$; we will call them 'Shepp's constants'. The aim of this chapter is to construct accurate approximations for $F_T(h)$ and hence for the Shepp's constants. We will demonstrate that at least some of the approximations are extremely accurate. The content of this chapter has been published in [81].

3.1 Introduction

Let $S(t)$, $t \in [0, T]$, be a Gaussian process with mean 0 and covariance

$$\mathbb{E}S(t)S(t') = \max\{0, 1 - |t - t'|\}. \quad (3.1.1)$$

This process is often called Slepian process. For any real h and $x < h$, define

$$F_T(h | x) := \Pr\left\{\max_{t \in [0, T]} S(t) < h \mid S(0) = x\right\}; \quad (3.1.2)$$

if $x \geq h$ we set $F_T(h | x) = 0$. Assuming that x has Gaussian distribution $N(0, 1)$, and hence the stationarity of the process $S(t)$, we average $F_T(h | x)$ and thus define

$$F_T(h) := \int_{-\infty}^h F_T(h | x) \varphi(x) dx, \quad (3.1.3)$$

where $\varphi(x) = (2\pi)^{-1/2} \exp\{-x^2/2\}$.

Key results on the boundary crossing probabilities for the Slepian process have been established by L. Shepp in [113]. In particular, Shepp has derived an explicit formula for $F_T(h)$ with T integer, see (3.2.5) below. As this explicit formula is quite complicated, in (3.7) in the same paper, Shepp has conjectured the existence of the following constant (depending on h)

$$\Lambda(h) = -\lim_{T \rightarrow \infty} \frac{1}{T} \log F_T(h) \quad (3.1.4)$$

and raised the question of constructing accurate approximations and bounds for this constant.

The importance of this constant is related to the asymptotic relation

$$F_T(h) \simeq \text{const}[\lambda(h)]^T \quad \text{as } T \rightarrow \infty, \quad (3.1.5)$$

where $\lambda(h) = \exp\{-\Lambda(h)\}$. We will call $\Lambda(h)$ and $\lambda(h)$ ‘Shepp’s constants’.

In this chapter, we are interested in deriving approximations for $F_T(h)$ in the form (3.1.5) and hence for the Shepp’s constants. In formulation of approximations, we offer approximations for $F_T(h)$ for all $T > 2$ and hence approximations for $\Lambda(h)$ and $\lambda(h)$. Note that computation of $F_T(h)$ for $T \leq 2$ is a relatively easy problem, see [122] for $T \leq 1$ and [113] for $1 < T \leq 2$.

In Section 3.2 we derive several approximations for $F_T(h)$ and $\lambda(h)$ and provide numerical results showing that at least some of the derived approximation are extremely accurate. In Section 3.3 we compare the upper tail asymptotics for the Slepian process and some other stationary Gaussian processes. Section 3.4 contains some minor technical details.

3.2 Construction of approximations

3.2.1 Existence of Shepp’s constants and the approximations derived from general principles

The fact that the limit in (3.1.4) exists and hence that $\Lambda(h)$ is properly defined for any h has been proven in [66]. The proof of existence of $\Lambda(h)$ is based on the inequalities

$$-\frac{1}{T+1} \log[F_T(h)] \leq \Lambda(h) \leq -\frac{1}{T} \log[F_T(h)] \quad \text{for any } T > 0; \quad (3.2.1)$$

see (3.5) in [66]. The upper estimate in (3.2.1) follows directly from the celebrated ‘Slepian inequality’ established in [123]; the Slepian inequality holds for any Gaussian stationary process whose correlation function is non-negative. The lower estimate in (3.2.1) can be obtained by a simple extension of the arguments in [123, p.470]; it holds for any Gaussian stationary process which correlation function vanishes outside the interval $[-1, 1]$. The inequalities (3.2.1) are not sharp: in particular, for $T = 2$ and $h = 0$, (3.2.1) gives $1.336 < \Lambda(0) < 2.004$; see [73, Remark 3]. From Tables 1 and 3, an accurate approximation for $\Lambda(0)$ is $\Lambda(0) \simeq 1.5972$, where the author claims all four decimal places are accurate.

If n is not too small, the bounds (3.2.1) are very difficult to compute. For small h , these bounds are not sharp even if n is large, see Figure 3.1a. The bounds improve as h grows, see Figure 3.1b. It is not very clear how to use these bounds for construction of accurate approximations for $\Lambda(h)$. In particular, from Figure 3.1b we observe that the upper bound of (3.2.1) can be much closer to the true $\Lambda(h)$ than the lower bound.

One may apply general results shown in [62, 69], see also formula (2.1.3) in [1], to approximate $F_T(h)$ for large h but these results only show that $\lambda(h) \rightarrow 1$ as $h \rightarrow \infty$ and therefore are of no use here. A more useful tool, which can be used for approximating $\lambda(h)$, is connected to the following result of J. Pickands proved in

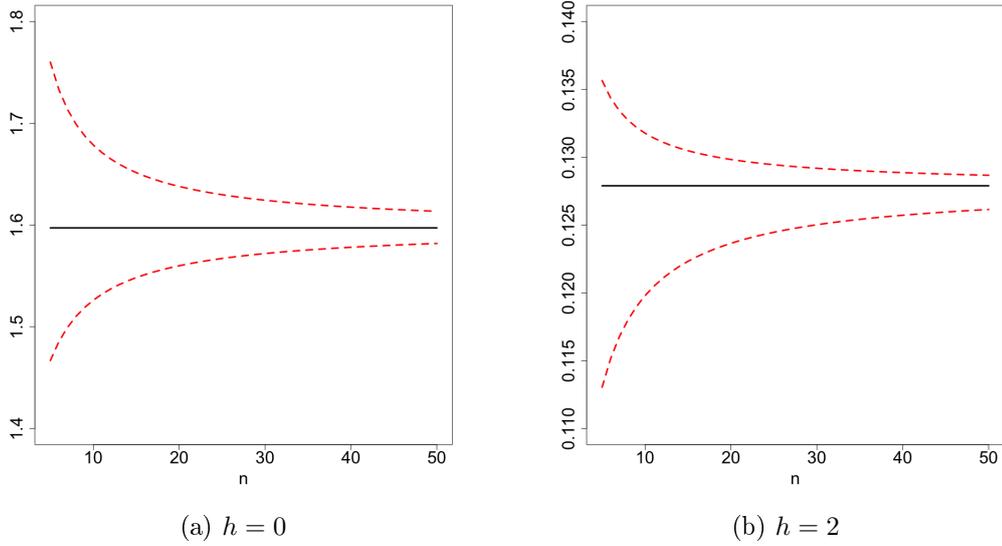


Figure 3.1: Lower and upper bounds (3.2.1) (red dotted lines) for $\Lambda(h)$ (solid black line).

[90]. Assume that $\{\xi(t)\}$ is a stationary Gaussian random process with $\mathbb{E}\xi(t) = 0$, $\mathbb{E}\xi^2(t) = 1$ and covariance function

$$\rho(t) = \mathbb{E}\xi(0)\xi(t) = 1 - C|t|^\alpha + o(|t|^\alpha) \quad \text{as } t \rightarrow 0 \quad (3.2.2)$$

and $\sup_{\epsilon \leq t \leq T} \rho(t) < 1, \forall \epsilon > 0$. Then

$$\Pr \left\{ \sup_{0 \leq t \leq T} \xi(t) \leq h \right\} = 1 - TC^{1/\alpha} H_\alpha h^{2/\alpha-1} \varphi(h) (1 + o(1)) \quad \text{as } h \rightarrow \infty, \quad (3.2.3)$$

where H_α is the so-called ‘Pickands constant’. By replacing $1 - x$ with e^{-x} ($x \rightarrow 0$) and removing the term $(1 + o(1))$ in (3.2.3) we obtain a general approximation

$$\Pr \left\{ \sup_{0 \leq t \leq T} \xi(t) \leq h \right\} \simeq \exp\{-TC^{1/\alpha} H_\alpha h^{2/\alpha-1} \varphi(h)\}. \quad (3.2.4)$$

As shown in [42], the value of the Pickands constant H_α is only known for $\alpha = 1, 2$ and hence the approximation (3.2.4) can only be applied in these cases. When $\xi(t)$ is the Slepian process $S(t)$ with covariance function (3.1.1) we have $\alpha = 1, H_1 = 1$ and $C = 1$. Hence we obtain from (3.2.4)

Approximation 0: $F_T(h) \simeq \exp(-h\varphi(h)T), \quad \Lambda^{(0)}(h) = h\varphi(h), \quad \lambda^{(0)}(h) = e^{-h\varphi(h)}.$

Note that Approximation 0 can also be obtained as a Poisson clumping heuristic, see formula (D10g) in [2]. If h is not large, then Approximation 0 is quite poor, see Tables 1 and 2 and Figure 2. For small and moderate values of h , the approximations derived below in this section are much superior to Approximation 0.

3.2.2 Shepp's formula for F_n

As previously discussed in Section 2.4.2 of Chapter 2, the following formula is the result (2.15) in [113]:

$$F_n(h|x) = \frac{1}{\varphi(x)} \int_{D_x} \det |\varphi(y_i - y_{j+1} + h)|_{i,j=0}^n dy_2 \dots dy_{n+1}, \quad (3.2.5)$$

where $T = n$ is a positive integer, $D_x = \{y_2, \dots, y_{T+1} | h-x < y_2 < y_3 < \dots < y_{n+1}\}$, $y_0 = 0, y_1 = h - x$. L. Shepp in [113] has also derived explicit formulas for $F_T(h|x)$ with non-integral $T > 0$ but these formulas are more complicated and are realistically applicable only for small T (say, $T \leq 3$).

From (3.2.5) we straightforwardly obtain

$$F_1(h|x) = \Phi(h) - \frac{\varphi(h)}{\varphi(x)} \Phi(x), \quad (3.2.6)$$

$$F_1(h) = \int_{-\infty}^h F_1(h|x) \varphi(x) dx = \Phi^2(h) - \varphi(h) [h\Phi(h) + \varphi(h)], \quad (3.2.7)$$

where $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$. Derivation of explicit formulas for $F_T(h|x)$ and $F_T(h)$ with $T \leq 1$ is relatively easy as the process $S(t)$ is conditionally Markovian in the interval $[0, 1]$, see [71]. Formula (3.2.6) has been first derived in [122].

In what follows, $F_2(h)$ also plays a very important role. Using (3.2.5) and changing the order of integration where suitable, $F_2(h)$ can be expressed through a one-dimensional integral as follows:

$$\begin{aligned} F_2(h) = & \Phi^3(h) - 2h\varphi(h)\Phi^2(h) + \frac{h^2 - 3 + \sqrt{\pi}h}{2} \varphi^2(h)\Phi(h) + \frac{h + \sqrt{\pi}}{2} \varphi^3(h) \\ & + \int_0^\infty \Phi(h-y) \left[\varphi(h+y)\Phi(h-y) - \sqrt{\pi}\varphi^2(h)\Phi(\sqrt{2}y) \right] dy. \end{aligned} \quad (3.2.8)$$

This expression can be approximated as shown in Appendix; see (3.4.1).

3.2.3 An alternative representation of the Shepp's formula (3.2.5)

From the discussions of Section 2.4.3 in Chapter 2, for $T = n$ a positive integer, $y_0 = 0, y_1 = h - x$ and (for $i = 0, 1, \dots, n$) by setting $s_i = h + y_i - y_{i+1}$ with $s_0 = x$, it follows from Shepp's proof of (3.2.5) that s_0, s_1, \dots, s_n have the meaning of the values of the process $S(t)$ at the times $t = 0, 1, \dots, n$: $S(i) = s_i$ ($i = 0, 1, \dots, n$). Changing the variables in (3.2.5), we obtain

$$F_n(h|x) = \frac{1}{\varphi(x)} \int_{-\infty}^h \dots \int_{-\infty}^h \det |\varphi(s_i + a_{i,j})|_{i,j=0}^n ds_1 \dots ds_n, \quad (3.2.9)$$

where

$$a_{i,j} = y_{i+1} - y_{j+1} = \begin{cases} 0 & \text{for } i = j \\ (i-j)h - s_{j+1} - \dots - s_{i+1} & \text{for } i > j \\ (i-j)h + s_{i+1} + \dots + s_j & \text{for } i < j. \end{cases}$$

Expression (3.2.9) for the probability $F_n(h|x)$ implies that the function

$$p(s_0, s_1, \dots, s_n) = \frac{1}{\varphi(s_0)F_n(h|s_0)} \det |\varphi(s_i + a_{i,j})|_{i,j=0}^n. \quad (3.2.10)$$

is the joint probability density function for the values $S(0), S(1), \dots, S(n)$ under the condition $S(t) < h$ for all $t \in [0, n]$.

Since s_n is the value of $S(n)$, the formula (3.2.10) also shows the transition density from $s_0 = x$ to s_n conditionally $S(t) < h$ for all $t \in [0, n]$:

$$p_h^{(n)}(x \rightarrow s_n) = \frac{1}{\varphi(x)} \int_{-\infty}^h \dots \int_{-\infty}^h \det |\varphi(s_i + a_{i,j})|_{i,j=0}^n ds_1 \dots ds_{n-1}. \quad (3.2.11)$$

For this transition density, $\int_{-\infty}^h p_h^{(n)}(x \rightarrow z) dz = F_n(h|x)$.

3.2.4 Approximating $\lambda(h)$ through eigenvalues of integral operators

3.2.4.1 One-step transition

In the case $n = 1$ we obtain from (3.2.11):

$$p_h^{(1)}(x \rightarrow z) = \frac{1}{\varphi(x)} \det \begin{pmatrix} \varphi(x) & \varphi(x-h+z) \\ \varphi(h) & \varphi(z) \end{pmatrix} = \varphi(z) [1 - e^{-(h-z)(h-x)}] \quad (3.2.12)$$

with $z = s_1 < h$. Let $\lambda_1(h)$ be the largest eigenvalue of the the integral operator with kernel (3.2.12):

$$\lambda_1(h)p(z) = \int_{-\infty}^h p(x)p_h^{(1)}(x \rightarrow z)dx, \quad z < h,$$

where eigenfunction $p(x)$ is some probability density on $(-\infty, h]$. The Ruelle-Krasnoselskii-Perron-Frobenius theory of bounded linear positive operators (see e.g. Theorem XIII.43 in [107]) implies that the maximum eigenvalue λ of the operator with kernel $K(x, z) = p_h^{(1)}(x \rightarrow z)$ is simple, real and positive and the eigenfunction $p(x)$ can be chosen as a probability density.

Similarly to what is done below in Section 3.2.4.2, we can suggest computing good numerical approximations to $\lambda_1(h)$ using Gauss-Legendre quadrature formulas. However, the use of (4.15) from [82] (a results not derived explicitly in this thesis) helps us to obtain the following simple but rather accurate approximation to $\lambda_1(h)$:

$$\hat{\lambda}_1(h) = \Phi(h) + \varphi(h)/h - \varphi(h)[\varphi(h) + h\Phi(h)]t / [\Phi(h) - e^{-h^2/2}/2].$$

Approximation 1: $F_T(h) \simeq F_1(h) [\hat{\lambda}_1(h)]^{T-1}$ ($T \geq 1$); $\Lambda^{(1)}(h) = -\log \hat{\lambda}_1(h)$,

$$\lambda^{(1)}(h) = \hat{\lambda}_1(h).$$

3.2.4.2 Transition in a twice longer interval

Consider now the interval $[0, 2]$. We could have extended the method of Section 3.2.4.1 and used the eigenvalue (square root of it) for the transition $s_0 \rightarrow s_2$ with transition density expressed in (3.2.11) with $n = 2$. This would improve Approximation 1 but this improvement is only marginal. Instead, we will use another approach: we consider the transition $s_1 \rightarrow s_2$ but use the interval $[0, 1]$ just for setting up the initial condition for observing $S(t)$ at $t \in [1, 2]$.

For $n = 2$, the expression (3.2.10) for the joint probability density function for the values $S(0), S(1), S(2)$ under the condition $S(t) < h$ for all $t \in [0, 2]$ has the form

$$p(s_0, s_1, s_2) = \frac{1}{\varphi(s_0)F_2(h|s_0)} \det \begin{pmatrix} \varphi(s_0) & \varphi(s_0 - h + s_1) & \varphi(s_0 - 2h + s_1 + s_2) \\ \varphi(h) & \varphi(s_1) & \varphi(s_1 + s_2 - h) \\ \varphi(2h - s_1) & \varphi(h) & \varphi(s_2) \end{pmatrix}.$$

Denote by $p_1(z)$, $z < h$, the ‘non-normalized’ density of $S(1)$ under the condition $S(t) < h$ for all $t \in [0, 1]$ that satisfies $\int_{-\infty}^h p_1(z) dz = F_1(h)$. Using (3.2.12), we obtain

$$p_1(z) = \int_{-\infty}^h p_h^{(1)}(x \rightarrow z) \varphi(x) dx = \Phi(h)\varphi(z) - \Phi(z)\varphi(h).$$

Then the transition density from $x = s_1$ to $z = s_2$ under the condition $S(t) < h$ for all $t \in [0, 2]$ is achieved by integrating s_0 out and renormalising the joint density:

$$\begin{aligned} q_h(x \rightarrow z) &= \frac{1}{p_1(x)} \int_{-\infty}^h \det \begin{pmatrix} \varphi(s_0) & \varphi(s_0 - h + x) & \varphi(s_0 - 2h + x + z) \\ \varphi(h) & \varphi(x) & \varphi(x + z - h) \\ \varphi(2h - x) & \varphi(h) & \varphi(z) \end{pmatrix} ds_0 \\ &= \frac{1}{\Phi(h)\varphi(x) - \Phi(x)\varphi(h)} \det \begin{pmatrix} \Phi(h) & \Phi(x) & \Phi(x + z - h) \\ \varphi(h) & \varphi(x) & \varphi(x + z - h) \\ \varphi(2h - x) & \varphi(h) & \varphi(z) \end{pmatrix}. \end{aligned}$$

Let $\lambda_2(h)$ be the largest eigenvalue of the integral operator with kernel q_h :

$$\lambda_2(h)q(z) = \int_{-\infty}^h q(x)q_h(x \rightarrow z)dx, \quad z < h,$$

where eigenfunction $q(x)$ is some probability density on $(-\infty, h]$. Similarly to the case $n = 1$, $\lambda_2(h)$ is simple, real and positive eigenvalue of the operator with kernel $K(x, z) = q_h(x \rightarrow z)$ and the eigenfunction $q(x)$ can be chosen as a probability density.

In numerical examples below we approximate $\lambda_2(h)$ with $\hat{\lambda}_2(h)$ using the methodology described in [72], p.154. We refer the reader to the end of Section 2.5.2.2 of Chapter 2 for a brief summary of this methodology.

Approximation 2: $F_T(h) \simeq F_2(h) \left[\hat{\lambda}_2(h) \right]^{T-2}$ ($T \geq 2$); $\Lambda^{(2)}(h) = -\log \hat{\lambda}_2(h)$,

$$\lambda^{(2)}(h) = \hat{\lambda}_2(h).$$

3.2.4.3 Quality of Approximations 1 and 2

Approximation 1 is more accurate than Approximation 0 but it is still not accurate enough. This is related to the fact that the process $S(t)$ is not Markovian and the behaviour of $S(t)$ on the interval $[i, i + 1]$ depends on all values of $S(t)$ in the interval $[i - 1, i]$ and not only on the value $s_i = S(i)$, which is a simplification we used for derivation of Approximation 1. Approximation 2 corrects the bias of Approximation 1 by considering twice longer intervals $[i - 1, i + 1]$ and using the behaviour of $S(t)$ in the first half of the interval $[i - 1, i + 1]$ just for setting up the initial condition at $[i, i + 1]$. As shown in Section 3.2.7, Approximation 2 is much more accurate than Approximations 0 and 1. The approximations developed in the following section also carefully consider the dependence of $S(t)$ on its past; they could be made arbitrarily accurate (on expense of increased computational complexity).

3.2.5 Main approximations

As mentioned above, the behaviour of $S(t)$ on the interval $[i, i + 1]$ depends on all values of $S(t)$ in the interval $[i - 1, i]$ and not only on the value $s_i = S(i)$. The exact value of the Shepp's constant $\lambda(h)$ can be defined as the limit (as $i \rightarrow \infty$) of the probability that $S(t) < h$ for all $t \in [i, i + 1]$ under the condition $S(t) < h$ for all $t \leq i$. Using the formula for conditional probability, we obtain

$$\lambda(h) = \lim_{i \rightarrow \infty} F_{i+1}(h)/F_i(h). \quad (3.2.13)$$

Waiting a long time without reaching h is not numerically possible and is not what is really required for computation of $\lambda(h)$. What we need is for the process $S(t)$ to (approximately) reach the stationary behaviour in the interval $[i - 1, i]$ under the condition $S(t) < h$ for all $t < i$. Since the memory of $S(t)$ is short (it follows from the representation $S(t) = W(t) - W(t + 1)$, where $W(t)$ is the standard Wiener process), this stationary behaviour of $S(t)$ is practically achieved for very small i , as is seen from numerical results of Section 3.2.7. Moreover, since ratios $F_{i+1}(h)/F_i(h)$ are very close to $F_{i+1}(h|x_h)/F_i(h|x_h)$ for $i \geq 1$, we can use ratios $F_{i+1}(h|x_h)/F_i(h|x_h)$ in (3.2.13) instead. Here $x_h = -\varphi(h)/\Phi(h)$ is the mean of the truncated normal distribution with density $\varphi(x)/\Phi(h)$, $x \leq h$. For computing the approximations, it makes integration easier. Note also another way of justifying the approximation $\lambda(h) \simeq F_{i+1}(h)/F_i(h)$: divide (3.1.5) with $T = i + 1$ by (3.1.5) with $T = i$.

The above considerations give rise to several approximations formulated below. We start with simpler approximations which are easy to compute and end up with approximations which are extremely accurate but are harder to compute. Approximation 7 is very precise, see Table 3.5. However, we would not recommend extremely accurate Approximations 6 and 7 since Approximations 4 and 5 are already very accurate, see Tables 1 and 2, but are much easier to compute. Approximation 3, the simplest in the family, is also quite accurate. Note that all approximations for $F_T(h)$ can be applied for any $T > 0$ (the accuracy of Approximations 4 and 7 for small T is shown in Table 3.3).

Approximation 3: $F_T(h) \simeq F_2(h) \left[\lambda^{(3)}(h) \right]^{T-2}$, where $\lambda^{(3)}(h) = F_2(h|x_h)/F_1(h|x_h)$.

Approximation 4: $F_T(h) \simeq F_2(h) \left[\lambda^{(4)}(h) \right]^{T-2}$, where $\lambda^{(4)}(h) = F_2(h)/F_1(h)$.

Approximation 5: $F_T(h) \simeq F_2(h) \left[\lambda^{(5)}(h) \right]^{T-2}$, where $\lambda^{(5)}(h) = F_3(h|x_h)/F_2(h|x_h)$.

Approximation 6: $F_T(h) \simeq F_3(h) \left[\lambda^{(6)}(h) \right]^{T-3}$, where $\lambda^{(6)}(h) = F_4(h|x_h)/F_3(h|x_h)$.

Approximation 7: $F_T(h) \simeq F_4(h) \left[\lambda^{(7)}(h) \right]^{T-4}$, where $\lambda^{(7)}(h) = F_4(h)/F_3(h)$.

Numerical complexity of these approximation is related to the necessity of computing either $F_n(h|0)$ or $F_n(h)$ for suitable n . It follows from (3.2.9) that $F_n(h|0)$ is an n -dimensional integral. Consequently, $F_n(h)$ is an $(n+1)$ -dimensional integral. In both cases, the dimensionality of the integral can be reduced by one, respectively to $n-1$ and n . In view of results of Sections 3.4.1 and 3.4.2, computation of Approximations 3 and 4 is easy, computation of Approximation 5 requires numerical evaluation of a one-dimensional integral (which is not hard) but to compute Approximation 7 we need to approximate a three-dimensional integral, which has to be done with high precision as otherwise Approximation 7 is not worth using: indeed, Approximations 4–6 are almost as good but are much easier to compute. As Approximation 7 provides us with the values which are practically indistinguishable from the true values of $\lambda(h)$, we use Approximation 7 only for the assessment of the accuracy of other approximations and do not recommend using it in practice.

3.2.6 Consistency of approximations when h is large

Assume that $h \rightarrow \infty$. We shall show that Approximations 3–7 for $\Lambda(h)$ give consistent results with Approximation 0 which is $\Lambda^{(0)}(h) = h\varphi(h)$.

Roughly, this consistency follows if we simply use $\Lambda^{(0)}(h)$ for $\Lambda(h)$ in (3.1.5) and then substitute the asymptotically correct values of $F_i(h)$ and $F_{i+1}(h)$ in $\Lambda(h) \simeq \log F_i(h) - \log F_{i+1}(h)$. Similar argument works in the case $\Lambda(h) \simeq \log F_i(h|x_h) - \log F_{i+1}(h|x_h)$.

Consider now Approximation 4 for $\Lambda(h)$, which is $\Lambda^{(4)}(h) = \log F_1(h) - \log F_2(h)$. From explicit formulas (3.2.7) and (3.2.8) for $F_1(h)$ and $F_2(h)$ we obtain

$$F_1(h) = 1 - \left(h + \frac{2}{h} + O\left(\frac{1}{h^3}\right) \right) \varphi(h), \quad F_2(h) = 1 - \left(2h - \frac{2}{h} + O\left(\frac{1}{h^2}\right) \right) \varphi(h), \quad h \rightarrow \infty. \quad (3.2.14)$$

Expansion of $F_1(h)$ is straightforward. To obtain the expansion of expansion of $F_2(h)$ from (3.2.8), we observe as $h \rightarrow \infty$:

$$\Phi^3(h) = 1 - \left(\frac{3}{h} + O\left(\frac{1}{h^3}\right) \right) \varphi(h), \quad 2h\varphi(h)\Phi^2(h) = \left(2h + O\left(\frac{1}{h^3}\right) \right) \varphi(h)$$

and

$$\int_0^\infty \Phi^2(h-y)\varphi(h+y)dy = \left(\frac{1}{h} + O\left(\frac{1}{h^3}\right) \right) \varphi(h);$$

all other terms in (3.2.8) converge to zero (as $h \rightarrow \infty$) faster than $\varphi(h)/h^2$. Using the expansion $\log(1-x) = -x + O(x^2)$ as $x \rightarrow 0$, this gives

$$\Lambda^{(4)}(h) = \log F_1(h) - \log F_2(h) = \left(h - \frac{4}{h} + O\left(\frac{1}{h^2}\right) \right) \varphi(h) \text{ as } h \rightarrow \infty.$$

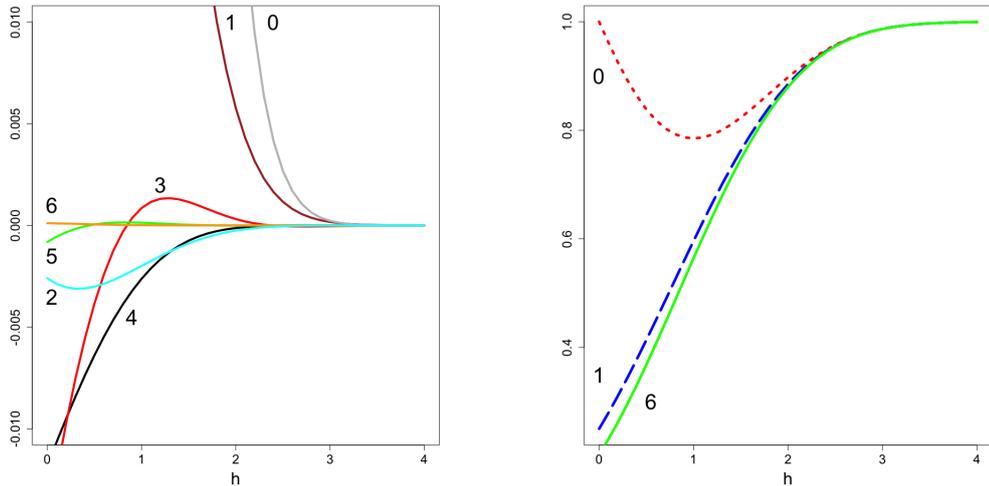
Similar considerations for the Approximation 5 give

$$\Lambda^{(5)}(h) = \log F_2(h|x_h) - \log F_3(h|x_h) = \left(h - \frac{3}{h} + O\left(\frac{1}{h^2}\right) \right) \varphi(h) \text{ as } h \rightarrow \infty.$$

This is fully consistent with approximation $\Lambda^{(0)}(h)$ and all the discussion of Section 3.2.1.

3.2.7 Numerical results

In this section we discuss the quality of approximations introduced in Section 3.2. In Table 3.1, we present the values of $\lambda^{(i)}(h)$, $i = 0, 1, \dots, 7$, for a number of different h ; see also Table 3.5 in Appendix at the end of this chapter. As mentioned above, $\lambda^{(7)}(h)$ is practically the true $\lambda(h)$ and therefore we compare all other approximations against $\lambda^{(7)}(h)$. A plot of the relative errors can be seen in Figure 3.2a, where the number next to the line corresponds to the approximation. Approximations 2 and 4 suggest accurate lower bounds for the true $\lambda(h)$. Approximations 0 and 1 appear to provide upper bounds for $\lambda(h)$ for all h . In Table 3.2 we present the relative errors of all other approximations against $\lambda^{(7)}(h)$; that is, the values $\lambda^{(i)}(h)/\lambda^{(7)}(h) - 1$ for $i = 0, 1, \dots, 6$. From these two tables we see that Approximations 2-6 are very accurate especially for $h > 1$.



(a) Relative errors of $\lambda^{(i)}(h)$, $i = 0, \dots, 6$, against $\lambda^{(7)}(h)$ (b) $\lambda^{(0)}(h)$ (dotted red), $\lambda^{(1)}(h)$ (dashed blue) and $\lambda^{(6)}(h)$ (solid green)

Figure 3.2: Approximations and their relative errors as functions of h .

As mentioned in Section 3.2.4.3, Approximation 1 is not as accurate as Approximations 2–7 because it does not adequately take into account the non-Markovianity

of $S(t)$. In Figure 3.2b we have plotted $\lambda^{(0)}(h)$ (dotted red line), $\lambda^{(1)}(h)$ (dashed red line) and $\lambda^{(6)}(h)$ (solid green line) for a range of interesting h . Visually, all $\lambda^{(i)}(h)$ with $i = 2, 4, 5, 6, 7$ would be visually indistinguishable from each other on the plot in Figure 3.2b and $\lambda^{(3)}(h)$ would be very close to them. The number next to the line corresponds to which approximation was used.

	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
$\lambda^{(0)}(h)$	1.000000	0.785079	0.897644	0.986792	0.999465
$\lambda^{(1)}(h)$	0.250054	0.596156	0.885025	0.986738	0.999466
$\lambda^{(2)}(h)$	0.201909	0.563246	0.879719	0.986566	0.999464
$\lambda^{(3)}(h)$	0.199421	0.564851	0.880220	0.986532	0.999463
$\lambda^{(4)}(h)$	0.200045	0.562888	0.879831	0.986570	0.999464
$\lambda^{(5)}(h)$	0.202269	0.564446	0.879943	0.986571	0.999464
$\lambda^{(6)}(h)$	0.202455	0.564377	0.879945	0.986571	0.999464
$\lambda^{(7)}(h)$	0.202434	0.564371	0.879945	0.986571	0.999464

Table 3.1: $\lambda^{(i)}(h), i = 0, 1, \dots, 7$, for different h .

	$h = 0$	$h=1$	$h=2$	$h=3$	$h=4$
$\lambda^{(0)}(h)$	3.94e+00	3.91e-01	2.01e-02	2.25e-04	6.12e-07
$\lambda^{(1)}(h)$	2.35e-01	5.63e-02	5.77e-03	1.69e-04	1.88e-06
$\lambda^{(2)}(h)$	-2.59e-03	-1.99e-03	-2.57e-04	-4.61e-06	-7.82e-09
$\lambda^{(3)}(h)$	-1.49e-02	8.51e-04	3.12e-04	-3.88e-05	-1.28e-06
$\lambda^{(4)}(h)$	-1.18e-02	-2.63e-03	-1.29e-04	-2.06e-07	1.35e-09
$\lambda^{(5)}(h)$	-8.13e-04	1.33e-04	-2.49e-06	-1.34e-07	9.09e-11
$\lambda^{(6)}(h)$	1.03e-04	1.09e-05	-1.83e-07	4.12e-11	6.09e-12

Table 3.2: Relative errors of $\lambda^{(i)}(h), i = 0, 1, \dots, 6$, against $\lambda^{(7)}(h)$.

In Table 3.3 we display values of $F_T(h)$ (numerically computed via (3.2.5)) and Approximations 4 and 7 for $h = 0, 1, 2$ and $T = 1, 2, 3, 4$. For larger values of T , a large-scale simulation study has been performed where $F_T(h)$ has been estimated for different h using 10^6 trajectories of $S(t)$ and all approximations for $F_T(h)$ considered above. Visually, Approximations 5-7 are virtually exact for all $h \geq 0$ and also Approximations 2-4 are visually undistinguishable from them for $h \geq 0.5$. In Figure 3.3, we show the strong performance of Approximation 7 for $F_T(h)$, where we have chosen $T = 10$ and $T = 50$.

Table 3.4 provides results of one run of a simulation experiment, where we have simulated the process $S(t)$. In this experiment (its results are very typical), we have chosen the interval $[0, 10)$ and repeated the simulation $N = 10^5$ times. To simulate the process $S(t)$, we have discretized the interval $[0, 10)$ into 10^4 points $j/1000$ ($j = 0, \dots, 10^4 - 1$) and run the moving window of size 1000. The value of the Shepp's constant $\lambda(2)$ corrected for discrete time is approximately 0.8851 (the uncorrected value of Shepp's constant $\lambda(2)$ is approximately 0.8799). Table 3.4 records values of N_i , the numbers of trajectories which have not reached the barrier $h = 2$ in $[0, i)$ (they were still running at time i), for $i = 1, \dots, 9$. $N_{10} = 28629$ trajectories, out of the

		$T=1$	$T=2$	$T=3$	$T=4$
$h = 0$	$F_T(h)$	0.090845	0.018173	0.003674	0.000744
	Approximation 4	0.090845	0.018173	0.003635	0.000727
	Approximation 7	0.089643	0.018149	0.003674	0.000744
$h = 1$	$F_T(h)$	0.445730	0.250896	0.141584	0.079906
	Approximation 4	0.445730	0.250896	0.141227	0.079495
	Approximation 7	0.444515	0.250871	0.141584	0.079906
$h = 2$	$F_T(h)$	0.846577	0.744845	0.655423	0.576737
	Approximation 4	0.846577	0.744845	0.655338	0.576587
	Approximation 7	0.846465	0.744844	0.655423	0.576737

Table 3.3: Values of $F_T(h)$ and Approximations 4 and 7 for different values of h and T .

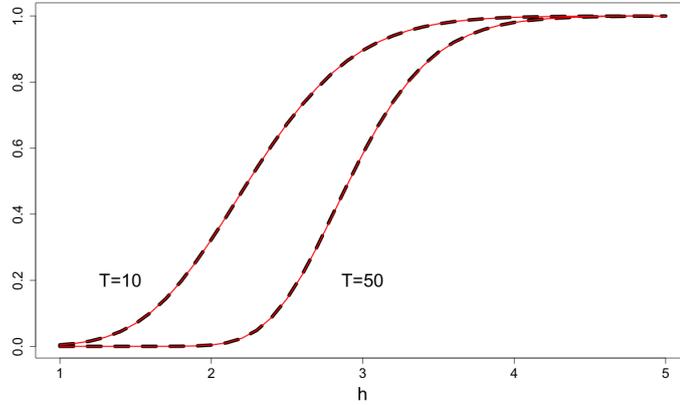


Figure 3.3: Empirical versions of $F_T(h)$ (dashed black) and Approximation 7 (solid red).

total number 10^5 of them, have not crossed the barrier in $[0, 10)$. The frequencies $f_i = N_{i+1}/N_i$ ($i = 1, \dots, 9$) are empirical versions of $F_{i+1}(h)/F_i(h)$ appearing in (3.2.13) with $h = 2$. The standard deviations associated with f_i 's are $s_i = \sqrt{f_i(1 - f_i)/N_i}$. The weighted empirical mean of $\{f_i\}$ is $\hat{f} = \sum_{i=1}^9 f_i N_i / \sum_{i=1}^9 N_i \simeq 0.8855$, which perfectly agrees with the Shepp's constant $\simeq 0.8851$. Both of these values lie well inside the $2s$ -confidence intervals $[f_i - 2s_i, f_i + 2s_i]$ for all $i \neq 8$ (for $i = 8$, these two values are very close to $f_i + 2s_i$). All numerical results fully support the considerations of Section 3.2.5.

i	1	2	3	4	5	6	7	8	9
N_i	85473	75732	67118	59405	52485	46516	41164	36367	32325
f_i	.8860	.8862	.8851	.8835	.8863	.8849	.8835	.8888	.8857
$2s_i$.0022	.0023	.0025	.0026	.0028	.0030	.0032	.0033	.0035

Table 3.4: Summary of results for an experiment involving $N = 10^5$ runs of $S(t)$ in $[0, 10)$ with $h = 2$.

3.3 Comparison of the upper tail asymptotics for the Slepian process against some other stationary Gaussian processes

Consider the following three stationary Gaussian processes.

1. $\xi_1(t)$ ($t \geq 0$) is the Ornstein-Uhlenbeck process with mean 0, variance 1 and correlation function $\rho_1(t) = \exp(-|t|)$.
2. Let $a > 0$ be fixed real number and set $\alpha = (1 + a + a^2)/(2 + 2a + a^2)$. Then, if $W(t)$ denotes the standard Wiener process, we define the process $\xi_2(t)$ ($t \geq 0$) as follows:

$$\xi_2(t) = \frac{1}{\sqrt{1 + a + a^2}} \{(1 + a)W(t + 2\alpha) - aW(t + \alpha) - W(t)\}.$$

The process $\xi_2(t)$ has mean 0, variance 1 and correlation function

$$\rho_2(t) = \begin{cases} 1 - |t|, & \text{for } 0 \leq |t| \leq \alpha \\ \frac{(1+a)(2\alpha-|t|)}{1+a+a^2}, & \text{for } \alpha \leq |t| \leq 2\alpha \\ 0 & \text{for } |t| \geq 2\alpha. \end{cases}$$

3. Let $c \geq 1$ be a fixed real number and set $\beta = 1/(c + 2)$. Define the process $\xi_3(t)$ by

$$\xi_3(t) = \frac{1}{\sqrt{1 + c^2}} \{W(t + 1) + cW(t + (c + 1)\beta) - cW(t + \beta) - W(t)\}.$$

The process $\xi_3(t)$ has mean 0, variance 1 and correlation function

$$\rho_3(t) = \begin{cases} 1 - |t| & \text{for } 0 \leq |t| \leq \beta \\ \frac{(1+c)(1+c^2\beta-|t|(1+c))}{1+c^2} & \text{for } \beta \leq |t| \leq c\beta \\ \frac{1+c+c^2\beta-|t|(1+2c)}{1+c^2} & \text{for } c\beta \leq |t| \leq (c + 1)\beta, \\ \frac{1-|t|}{1+c^2} & \text{for } (c + 1)\beta \leq |t| \leq 1 \\ 0 & \text{for } |t| \geq 1. \end{cases}$$

It follows from [71, Theorem 3], that the above three processes provide a very good representation of the entire class of conditionally Markov stationary Gaussian processes. Indeed, there is only one process in this class where $\alpha \neq 1$ in (3.2.2) (this is the process with covariance function $\rho(t) = \cos \omega t$ with $\omega \neq 0$) and the three types of processes we consider cover well the case where $\alpha = 1$ and $C = 1$ in (3.2.2) (the case $C \neq 1$ reduces to the case $C = 1$ by substituting h/C for h). For a graphical representation of the chosen covariance functions, see Figure 3.4b.

Below we compare Shepp's constant $\Lambda(h)$ defined in (3.1.4) to similar quantities of the processes $\{\xi_i(t)\}$, ($i = 1, 2, 3$) defined above. More precisely, let

$$F_{T,i}(h) := \Pr \left\{ \max_{t \in [0, T]} \xi_i(t) < h \right\}, \quad i = 1, 2, 3.$$

We are interested in comparing Shepp's constant $\Lambda(h)$ with

$$\Lambda_i(h) = - \lim_{T \rightarrow \infty} \frac{1}{T} \log F_{T,i}(h), \quad (3.3.1)$$

for $i = 1, 2, 3$. Importantly, each process has $\mathbb{E}\xi_i(t) = 0$, $\mathbb{E}\xi_i^2(t) = 1$ and correlation function $\rho_i(t) = \mathbb{E}\xi_i(0)\xi_i(t)$ which satisfies $\rho'_i(0^+) = \frac{d}{dt}\rho_i(t)|_{t=0^+} = -1$, for $i = 1, 2, 3$.

The existence and evaluation of the constant $\Lambda_1(h)$ defined in (3.3.1) for the Ornstein-Uhlenbeck process has been considered in [7], where it was shown that $0 < \Lambda_1(h) < 1$ for all $h > 0$, and that $\Lambda_1(h)$ is the root of a parabolic cylinder function (defined in [7]) closest to zero. It is also shown that $\lim_{h \rightarrow 0^+} \Lambda_1(h) = 1$ and $\lim_{h \rightarrow \infty} \Lambda_1(h) = h\varphi(h)$. The existence of the constants $\Lambda_2(h)$ and $\Lambda_3(h)$ follows from similar arguments for the existence of Shepp's constant $\Lambda(h)$. Moreover, the constants are approximated by the same methodology as Shepp's constant, namely:

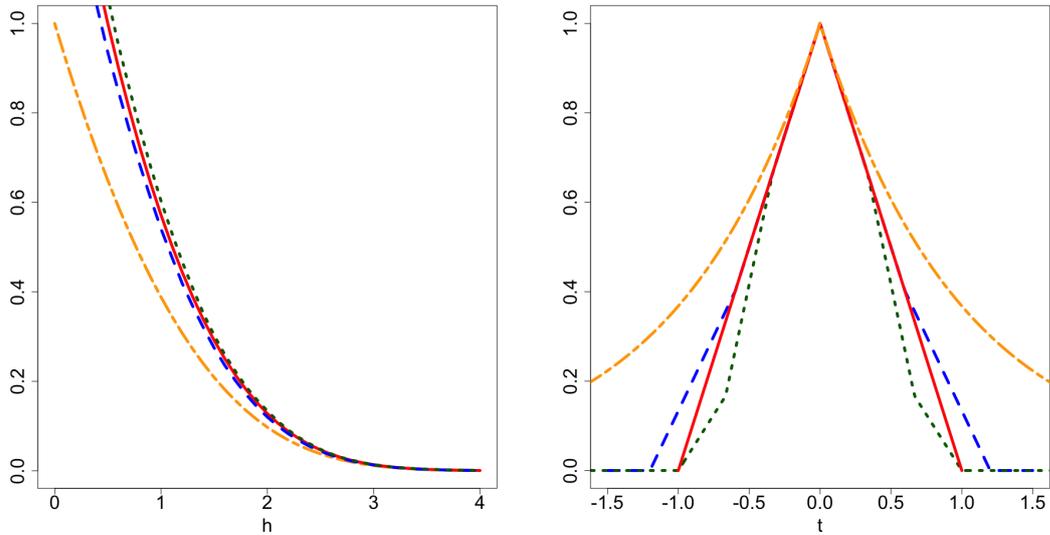
$$\lambda_i(h) = \exp(-\Lambda_i(h)) = \lim_{j \rightarrow \infty} F_{j+1,i}(h)/F_{j,i}(h). \quad (3.3.2)$$

The justification why we expect $F_{j+1,i}(h)/F_{j,i}(h)$ (with, say, $j \geq 3$) to be a good approximation of $\lambda_i(h)$ is related to the property of 'fast loss of memory', which processes $\xi_2(t)$ and $\xi_3(t)$ possess, as the process $S(t)$ does. In view of the complex structure of $\xi_2(t)$ and $\xi_3(t)$, the values of $F_{T,2}(h)$ and $F_{T,3}(h)$ are evaluated via Monte Carlo simulations. In Figure 3.4a, we compare $\Lambda_i(h)$ with Shepp's constant $\Lambda(h)$ (red solid line). $\Lambda_1(h)$ (orange dot-dash line) has been computed as in [7]. $\Lambda_2(h)$ (blue dashed line) and $\Lambda_3(h)$ (dark green dotted line) have been approximated using (3.3.2) with $j = 3$. For $\Lambda_2(h)$ we have taken $a = 1$ in the definition of $\xi_2(t)$ and for $\Lambda_3(h)$ we have taken $c = 1$ in the definition of $\xi_3(t)$. In Figure 3.4b, we plot the correlation functions: $\rho(t)$ (red solid line); $\rho_1(t)$ (orange dot-dash line); $\rho_2(t)$ (blue dashed line); $\rho_3(t)$ (dark green dotted line). Note that the results obtained are fully consistent with the celebrated 'Slepian's lemma', a Gaussian comparison inequality, see Lemma 1 in [123]. In our terms, Slepian's lemma says that if for two stationary Gaussian processes with non-negative covariance functions ρ_1 and ρ_2 we have $\rho_1(t) \geq \rho_2(t)$ for all $t \geq 0$, then for the corresponding values of $\Lambda(h)$ we have $\Lambda_1(h) \leq \Lambda_2(h)$, for all h .

3.4 Appendix

3.4.1 Approximations for $\Lambda(h)$

In Table 3.5 we use Approximation 7 (the most accurate approximation provided in this chapter) to approximate $\Lambda(h)$ over increments 0.1 for h . Bold font indicates the decimal places which the author claims as accurate. Note that $h = 0$ has been treated as a special case, see for example [123] and [91]. For $h = 0$, instead of Approximation 7, the approximation $\Lambda^{(8)}(h) = -\log(F_5(h)/F_4(h))$ has been used; the author does not recommend using this approximation in general because of its high complexity.



(a) Shepp's constant $\Lambda(h)$ and $\Lambda_i(h)$, $i = 1, 2, 3$ (b) Correlation functions $\rho(t)$ and $\rho_i(t)$, $i = 1, 2, 3$

Figure 3.4: Comparison of the upper tail asymptotics for several Gaussian stationary process

h	$\Lambda(h)$	h	$\Lambda(h)$	h	$\Lambda(h)$	h	$\Lambda(h)$	h	$\Lambda(h)$
0.0	1.5972	0.8	0.7240	1.6	0.250519	2.4	0.0578944	3.2	0.0077016
0.1	1.4632	0.9	0.6450	1.7	0.213929	2.5	0.0464986	3.3	0.0057244
0.2	1.3365	1.0	0.5720	1.8	0.181484	2.6	0.0370122	3.4	0.0042111
0.3	1.2170	1.1	0.5051	1.9	0.152902	2.7	0.0291909	3.5	0.0030658
0.4	1.1047	1.2	0.4438	2.0	0.127896	2.8	0.0228058	3.6	0.0022087
0.5	0.9995	1.3	0.3879	2.1	0.106178	2.9	0.0176462	3.7	0.0015747
0.6	0.9010	1.4	0.3372	2.2	0.087460	3.0	0.0135203	3.8	0.0011109
0.7	0.8092	1.5	0.2915	2.3	0.071458	3.1	0.0102561	3.9	0.0007755

Table 3.5: Approximations for $\Lambda(h)$ with accurate decimal digits in bold.

3.4.2 An approximation for $F_2(h)$

Using approximations for $\Phi(t)$, it is possible to accurately approximate the one-dimensional integral

$$I = \int_0^\infty \Phi(h-y) \left[\varphi(h+y)\Phi(h-y) - \sqrt{\pi}\varphi^2(h)\Phi(\sqrt{2}y) \right] dy$$

from the formula (3.2.8) for $F_2(h)$. For example, using the approximation (see [67])

$$\Phi(t) = \begin{cases} 0.5 \exp(0.717t - 0.416t^2) & \text{for } t \leq 0 \\ 1 - 0.5 \exp(-0.717t - 0.416t^2) & \text{for } t > 0, \end{cases}$$

we obtain the following approximation:

$$I \simeq \hat{I} = \Phi(2h) - \Phi(h) - \frac{\sqrt{\pi}\varphi^2(h)C_1}{2} [U(0.416, 0.832h - 0.717) - U(1.248, 0.832h - 1.731)/2] \\ - \frac{1}{4\sqrt{2\pi}} [4C_2J(0.916, 0.717 - 0.168h) - C_3J(1.332, 1.434 + 0.664h) - C_4U(1.332, 0.664h - 1.434)] \\ - \sqrt{\pi}\varphi^2(h) \left[h - J(0.832, -1.014)/2 - \frac{C_5}{2} \{J(0.416, 0.717 + 0.832h) - J(1.248, -0.297 + 0.832h)/2\} \right]$$

where $C_1 = e^{0.717h - 0.416h^2}$, $C_2 = e^{-0.717h - 0.916h^2}$, $C_3 = e^{-1.434h - 1.332h^2}$, $C_4 = e^{1.434h - 1.332h^2}$, $C_5 = e^{-0.717h - 0.416h^2}$, $K(x, y, z) = \sqrt{\pi}x^{-1/2}e^{y^2/(4x)}\Phi(\sqrt{2x}z - y/\sqrt{2x})$, $J(x, y) = K(x, y, h) - K(x, y, 0)$ and $U(x, y) = K(x, y, \infty) - K(x, y, h)$.

Table 3.6 shows that \hat{I} is a rather accurate approximation for I across all h of interest.

	$h=0$	$h=0.5$	$h=1$	$h=1.5$	$h=2$	$h=2.5$	$h=3$	$h=3.5$
I	-0.043731	-0.046973	-0.016129	0.009474	0.011033	0.004760	0.001242	0.000228
\hat{I}	-0.043358	-0.047742	-0.016829	0.009355	0.011041	0.004763	0.001243	0.000228

Table 3.6: Values of I and its approximation \hat{I} for different values of h .

Chapter 4

First passage times for Slepian process with linear and piecewise linear barriers

Abstract

In this chapter, we derive explicit formulas for the first-passage probabilities of the process $S(t) = W(t) - W(t+1)$, where $W(t)$ is the Brownian motion, for linear and piece-wise linear barriers on arbitrary intervals $[0, T]$. Previously, explicit formulas for the first-passage probabilities of this process were known only for the cases of a constant barrier or $T \leq 1$. The first-passage probabilities results are used to derive explicit formulas for the power of a familiar test for change-point detection in the Wiener process. They will be useful when we consider the discrete-time analogue change-point problem in Chapter 5. The content of this chapter has been published in [145].

4.1 Introduction

Let $T > 0$ be a fixed real number and let $S(t)$, $t \in [0, T]$, be a Gaussian process with mean 0 and covariance

$$\mathbb{E}S(t)S(t') = \max\{0, 1 - |t - t'|\}.$$

This process is often called Slepian process and can be expressed in terms of the standard Brownian motion $W(t)$ by

$$S(t) = W(t) - W(t+1), \quad t \geq 0. \quad (4.1.1)$$

Let a and b be fixed real numbers and $x < a$. We are interested in an explicit formula for the first-passage probability

$$F_{a,b}(T | x) := \Pr(S(t) < a + bt \text{ for all } t \in [0, T] \mid S(0) = x); \quad (4.1.2)$$

note $F_{a,b}(T | x) = 0$ for $x \geq a$.

The case of a constant barrier, when $b = 0$, has attracted significant attention in literature. In his seminal paper [122], D. Slepian has shown how to derive an

explicit expression for $F_{a,0}(T|x)$ in the case $T \leq 1$; see also [71]. The case $T > 1$ is much more complicated than the case $T \leq 1$. Explicit formulas for $F_{a,0}(T|x)$ with general T were derived by L. A. Shepp in [113]; these formulas are special cases of results formulated in Section 4.2 and have been discussed previously in Chapter's 2 and 3. This chapter can be considered as a natural extension of the methodology developed in [113] and [122]; hence the title of this chapter. Expressions of the first passage probability $F_{a,b}(T|x)$ (and extensions including piecewise linear barriers) are necessary for computing the power of the MOSUM test when applied to the change-point problem of identifying a transient change in the mean of a Wiener process; see Section 4.5. The results of this chapter are crucial for Chapter 5, where the discrete time analogue change-point problem of identifying a transient change in mean of i.i.d. normal random variables is considered.

In the case $T \leq 1$, Slepian's method for deriving formulas for $F_{a,0}(T|x)$ can be easily extended to the case of a general linear barrier. An explicit formula for the first-passage probability $F_{a,b}(T|x)$ was first derived in 1988 in [143, p.81] (published in Russian) and more than 20 years later it was independently derived in [11] and [21]. In [143], the first-passage probability $F_{a,b}(T|x)$ for $T \leq 1$ was obtained by using the fact that $S(t)$ is a conditionally Markov process on the interval $[0, 1]$. It was shown in [71], and discussed in Section 2.3.2 of Chapter 2, that after conditioning on $S(0) = x$, $S(t)$ can be expressed in terms of the Brownian motion by $S(t) = (2 - t)W(g(t)) + x(1 - t)$ ($0 \leq t \leq 1$) with $g(t) = t/(2 - t)$. Consequently, the first-passage probabilities for $S(t)$, $t \in [0, T]$ with $T \leq 1$ can be obtained using first-passage formulas for the Brownian motion. This methodology, like many others, fails for $T > 1$.

For general $T > 0$, including the case $T > 1$, explicit formulas for $F_{a,b}(T|x)$ were unknown. Derivation of these formulas is the main objective of this chapter. To do this, we generalise Shepp's methodology of [113]. The principal distinction between Shepp's methodology and the results of this chapter is the use of an alternative way of computing coincidence probabilities. Shepp's proofs heavily rely on the so-called Karlin-McGregor identity, see [49]; we use an extension of this identity formulated in [50] and discussed in Section 4.2.1.

The Karlin-McGregor identity has many deep implications in probability. In [50, 54], the identity was used to show a connection between n independent Brownian motion processes conditioned to never collide and eigenvalues of random matrices. More specifically, if $X(t)$ represents a system of n independent Brownian motions starting from the origin and conditioned never to collide with each other, then the distribution of $X(t)$ can be obtained using the probability density of eigenvalues of random matrices in the Gaussian Unitary Ensemble, also see [52, 55]. Moreover, if an appropriate initial distribution of $X(t)$ is used, then it can be shown that non-colliding Brownian motion is a determinantal process; by this, we mean that any joint transition density can be expressed by a determinant of a matrix kernel, see [53]. In [13], after a slight generalisation of the Karlin-McGregor identity (a generalisation different to the one used in this chapter), the authors show applications in queuing theory. Another important application of the Karlin-McGregor identity deals with finding boundary crossing probabilities for various scan statistics, see [33, 78, 83].

The structure of this chapter is as follows. In Section 4.2.2, we provide an expression for $F_{a,b}(T|x)$ for integer T and in Section 4.2.4 we extend the results for non-integer T . In Sections 4.3 and 4.4, we extend the results to the case of piecewise-linear barriers. In Section 4.5, we outline an application to a change-point detection problem; this application was the authors main motivation for this research. In the appendix located at the end of the chapter, we provide detailed proofs of all theorems.

4.2 Linear barrier $a + bt$

The key result of this section is Theorem 4.2.1, where an explicit formula is derived for the first-passage probability $F_{a,b}(T|x)$ defined in (4.1.2) under the assumption that T is a positive integer, $T = n$. First, we formulate a lemma that is key to the advances of this chapter and can be obtained from [50, p. 5] or [51, p.40]. In this lemma, we use the notation

$$\varphi_s(z) := \frac{1}{\sqrt{2\pi s}} e^{-z^2/(2s)} \quad (4.2.1)$$

for the normal density with variance s . For the standard Brownian motion process $W(t)$, $\varphi_s(a - c)dc = \Pr(W(s) \in dc | W(0) = a)$ is the transition probability. We shall also use

$$\mathbb{W}_{n+1} = \{\mathbf{x} = (x_0, \dots, x_n)' \in \mathbb{R}^{n+1} : x_0 < x_1 < \dots < x_n\}$$

for the so-called Weyl chamber of type A_n , see [27] for details. The Weyl chamber of type A_n has a natural appearance when considering the first passage probability $F_{a,b}(T|x)$ as we shall highlight here; see the proof of Theorem 4.2.2 for exact details. For example, if one considers the probability $F_{a,b}(1|x)$, then using (4.1.1) we have

$$F_{a,b}(1|x) = \Pr(W(t) < a + bt + W(t+1) \text{ for all } t \in [0, 1] | S(0) = x).$$

From this, it is clear we are interested in the probability one Brownian motion process never collides with another slightly shifted Brownian motion process (here for brevity, we are disregarding the fact the Brownian motions are correlated). As a result, we are interested in the probability that two Brownian motion processes remain in the Weyl chamber of type A_n .

4.2.1 An important auxiliary result

Lemma 4.2.1 (From [50, p. 5]) *For any $s > 0$ and a positive integer n , let $\mathbf{W}^\mu(t) := (W_0(t), W_1(t), \dots, W_n(t))$, $t \in [0, s]$, be an $(n+1)$ -dimensional Brownian motion process with drift $\boldsymbol{\mu} = (\mu_0, \mu_1, \dots, \mu_n)'$. Then*

$$\begin{aligned} & \Pr\{\mathbf{W}^\mu(t) \in \mathbb{W}_{n+1} \ \forall t \in [0, s], \mathbf{W}^\mu(s) \in d\mathbf{c} \mid \mathbf{W}^\mu(0) = \mathbf{a}\} \\ &= \exp\left(-\frac{s}{2}\|\boldsymbol{\mu}\|^2 + \boldsymbol{\mu}'(\mathbf{c} - \mathbf{a})\right) \det[\varphi_s(a_i - c_j)]_{i,j=0}^n dc_0 dc_1 \dots dc_n \end{aligned} \quad (4.2.2)$$

where $\|\cdot\|$ denotes the Euclidean norm, $\mathbf{a} = (a_0, a_1, \dots, a_n)' \in \mathbb{W}_{n+1}$, $\mathbf{c} = (c_0, c_1, \dots, c_n)' \in \mathbb{W}_{n+1}$ and $d\mathbf{c} = (dc_0, \dots, dc_n)$, where dc_0, \dots, dc_n are infinitesimal intervals around c_0, \dots, c_n .

Lemma 4.2.1 is an extension of the Karlin-McGregor identity of [49], when applied specifically to the Brownian motion, and accommodates for different drift parameters μ_i of $W_i(t)$.

Corollary 4.2.1 *Under the same assumptions as Lemma 4.2.1, we have*

$$\Pr\{\mathbf{W}^\mu(t) \in \mathbb{W}_{n+1} \forall t \in [0, s] \mid \mathbf{W}^\mu(0) = \mathbf{a}, \mathbf{W}^\mu(s) = \mathbf{c}\} \\ = \exp\left(-\frac{s}{2}\|\boldsymbol{\mu}\|^2 + \boldsymbol{\mu}'(\mathbf{c} - \mathbf{a})\right) \det[\varphi_s(a_i - c_j)]_{i,j=0}^n / \prod_{i=0}^n \varphi_s(a_i - c_i + \mu_i s). \quad (4.2.3)$$

Proof. Denote the transition density for the process $W_i(t)$ by $\varphi_{s,\mu_i}(a - c)$; that is, $\varphi_{s,\mu_i}(a - c)dc = \Pr(W_i(s) \in dc \mid W_i(0) = a)$. Using the relation $\varphi_{s,\mu_i}(a - c) = \varphi_s(a - c + \mu_i s)$ and dividing both sides of (4.2.2) by $\Pr(\mathbf{W}^\mu(s) \in d\mathbf{c} \mid \mathbf{W}^\mu(0) = \mathbf{a})$, we obtain the result. \square

4.2.2 Linear barrier $a + bt$ with integer T

Let $\varphi(t) = \varphi_1(t)$ and $\Phi(t) = \int_{-\infty}^t \varphi(u)du$ be the density and the c.d.f. of the standard normal distribution. Assume that $T = n$ is a positive integer. Define $(n+1)$ -dimensional vectors

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ b \\ 2b \\ \vdots \\ nb \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ x_1 + a \\ x_2 + 2a + b \\ \vdots \\ x_n + na + \frac{(n-1)n}{2}b \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} x_1 \\ x_2 + a + b \\ x_3 + 2a + 3b \\ \vdots \\ x_{n+1} + (a+b)n + \frac{(n-1)n}{2}b \end{bmatrix} \quad (4.2.4)$$

and let μ_i , a_i and c_i be i -th components of vectors $\boldsymbol{\mu}$, \mathbf{a} and \mathbf{c} respectively ($i = 0, 1, \dots, n$). Note that we start the indexation of vector components at 0.

Theorem 4.2.1 *For any integer $n \geq 1$ and $x < a$,*

$$F_{a,b}(n \mid x) = \frac{1}{\varphi(x)} \int_{-x-a-b}^{\infty} \int_{x_2-a-2b}^{\infty} \cdots \int_{x_n-a-nb}^{\infty} \exp(-\|\boldsymbol{\mu}\|^2/2 + \boldsymbol{\mu}'(\mathbf{c} - \mathbf{a})) \\ \times \det[\varphi(a_i - c_j)]_{i,j=0}^n dx_{n+1} dx_n \dots dx_2, \quad (4.2.5)$$

where $\boldsymbol{\mu}$, \mathbf{a} and \mathbf{c} are given in (4.2.4).

Theorem 4.2.1 is a special case of Theorem 4.3.1 with (using the notation of Theorem 4.3.1) $n = T$ and $T' = 0$. Theorem 4.2.1 is formulated as a separate theorem as it is the first natural extension of Shepp's results of [113]. Indeed, if $b = 0$ then $\boldsymbol{\mu} = 0$ and (4.2.5) coincides with Shepp's formula (2.15) in [113] expressed in the variables $y_i = x_i + ia$ ($i = 0, 1, \dots, n$).

4.2.3 An alternative representation of formula (4.2.5)

It is easier to interpret Theorem 4.2.1 by expressing the integrals in terms of the values of $S(t)$ at times $t = 0, 1, \dots, n$. Let $x_0 = 0, x_1 = -x$. For $i = 0, 1, \dots, n$ we

set $s_i = x_i - x_{i+1}$ with $s_0 = x$. It follows from the proof of (4.2.5), see Section 4.6.2, that s_0, s_1, \dots, s_n have the meaning of the values of the process $S(t)$ at times $t = 0, 1, \dots, n$; that is, $S(i) = s_i$ ($i = 0, 1, \dots, n$). The range of the variables s_i in (4.2.5) is $(-\infty, a + bi)$, for $i = 0, 1, \dots, n$. The variables x_1, \dots, x_{n+1} are expressed via s_0, \dots, s_n by $x_k = -s_0 - s_1 - \dots - s_{k-1}$ ($k = 1, \dots, n+1$) with $x_0 = 0$. Changing the variables, we obtain the following equivalent expression for the probability $F_{a,b}(n | x)$:

$$F_{a,b}(n | x) = \frac{1}{\varphi(x)} \int_{-\infty}^{a+b} \int_{-\infty}^{a+2b} \dots \int_{-\infty}^{a+bn} \exp(-\|\boldsymbol{\mu}\|^2/2 + \boldsymbol{\mu}'(\mathbf{c} - \mathbf{a})) \\ \times \det [\varphi(a_i - c_j)]_{i,j=0}^n ds_n \dots ds_2 ds_1,$$

where $\boldsymbol{\mu}$ is given by (4.2.4) but expressions for \mathbf{a} and \mathbf{c} change:

$$\mathbf{a} = \begin{bmatrix} 0 \\ a - s_0 \\ 2a + b - s_0 - s_1 \\ \vdots \\ na + \frac{(n-1)n}{2}b - s_0 - s_1 - \dots - s_{n-1} \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -s_0 \\ a + b - s_0 - s_1 \\ 2a + 3b - s_0 - s_1 - s_2 \\ \vdots \\ (a+b)n + \frac{(n-1)n}{2}b - s_0 - s_1 - \dots - s_n \end{bmatrix}.$$

In a particular case of $n = 1$ we obtain:

$$F_{a,b}(1 | x) = \frac{1}{\varphi(x)} \int_{-\infty}^{a+b} \exp(-b^2/2 + b(b - s_1)) \det \begin{bmatrix} \varphi(x) & \varphi(x + s_1 - a - b) \\ \varphi(a) & \varphi(s_1 - b) \end{bmatrix} ds_1 \\ = \Phi(a + b) - \exp(-(a^2 - x^2)/2 - b(a - x)) \Phi(x + b), \quad (4.2.6)$$

which agrees with results in [11, 21, 143].

4.2.4 Linear barrier $a + bt$ with non-integer T

In this section, we shall provide an explicit formula for the first-passage probability $F_{a,b}(T | x)$ defined in (4.1.2) assuming $T > 0$ is not an integer. Represent T as $T = m + \theta$, where $m = [T] \geq 0$ is the integer part of T and $0 < \theta < 1$. Set $n = m + 1 = [T]$.

Let $\varphi_\theta(t)$ and $\varphi_{1-\theta}(t)$ be as defined in (4.2.1). Define the $(n+1)$ - and n -dimensional vectors as follows: $\boldsymbol{\mu}_1 = \boldsymbol{\mu}$ is as defined in (4.2.4),

$$\mathbf{a}_1 = \begin{bmatrix} 0 \\ u_1 + a \\ u_2 + 2a + b \\ \vdots \\ u_n + na + \frac{n(n-1)}{2}b \end{bmatrix}, \quad \mathbf{c}_1 = \begin{bmatrix} v_0 \\ v_1 + a + b\theta \\ v_2 + 2(a + b\theta) + b \\ \vdots \\ v_n + n(a + b\theta) + \frac{n(n-1)}{2}b \end{bmatrix}, \quad (4.2.7)$$

$$\boldsymbol{\mu}_2 = \begin{bmatrix} 0 \\ b \\ 2b \\ \vdots \\ mb \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} v_0 \\ v_1 + a + b\theta \\ v_2 + 2(a + b\theta) + b \\ \vdots \\ v_m + m(a + b\theta) + \frac{(m-1)m}{2}b \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} u_1 \\ u_2 + a + b \\ u_3 + 2a + 3b \\ \vdots \\ u_{m+1} + m(a + b) + \frac{(m-1)m}{2}b \end{bmatrix}, \quad (4.2.8)$$

and let a_{1i} and c_{1i} be i -th components of vectors \mathbf{a}_1 and \mathbf{c}_1 respectively ($i = 0, 1, \dots, n$). Similarly, let a_{2i} and c_{2i} be i -th components of vectors \mathbf{a}_2 and \mathbf{c}_2 respectively ($i = 0, 1, \dots, m$). Recall that we start the indexation of vector components at 0.

Theorem 4.2.2 *For $x < a$ and non-integer $T = m + \theta$ with $0 < \theta < 1$, we have*

$$F_{a,b}(T|x) = \frac{1}{\varphi(x)} \int_{-x-a-b}^{\infty} \cdots \int_{u_m-a-mb}^{\infty} \int_{-\infty}^{\infty} \int_{v_0-a-b\theta}^{\infty} \cdots \int_{v_m-a-b\theta-mb}^{\infty} \exp(-\theta\|\boldsymbol{\mu}_1\|^2/2 + \boldsymbol{\mu}'_1(\mathbf{c}_1 - \mathbf{a}_1)) \exp(-(1-\theta)\|\boldsymbol{\mu}_2\|^2/2 + \boldsymbol{\mu}'_2(\mathbf{c}_2 - \mathbf{a}_2)) \times \det[\varphi_\theta(a_{1i} - c_{1j})]_{i,j=0}^n \det[\varphi_{1-\theta}(a_{2i} - c_{2j})]_{i,j=0}^m dv_{m+1} \dots dv_1 dv_0 du_{m+1} \dots du_2.$$

A proof of Theorem 4.2.2 is provided in Section 4.6.1. If $b = 0$ then the above formula for $F_{a,b}(T|x)$ coincides with Shepp's formula (2.25) in [113] expressed in variables $x_i = u_i + ia$ and $y_i = v_i + ia$ ($i = 0, 1, \dots, n$). For $m = 0$ and hence $T = \theta$, Theorem 4.2.2 agrees with results in [11, 21, 143].

4.3 Piecewise linear barrier with one change of slope

4.3.1 Boundary crossing probability

In this section, we provide an explicit formula for the first-passage probability for $S(t)$ with a continuous piecewise linear barrier, where not more than one change of slope is allowed. For any non-negative T, T' and real a, b, b' we define the piecewise-linear barrier $B_{T,T'}(t; a, b, b')$ by

$$B_{T,T'}(t; a, b, b') = \begin{cases} a + bt & t \in [0, T], \\ a + bT + b'(t - T) & t \in [T, T + T']; \end{cases}$$

for an illustration of this barrier, see Figure 4.1. We are interested in finding an expression for the first-passage probability

$$F_{a,b,b'}(T, T' | x) := \Pr(S(t) < B_{T,T'}(t; a, b, b') \text{ for all } t \in [0, T + T'] | S(0) = x). \quad (4.3.1)$$

We only consider the case when both T and T' are integers. The case of general T, T' can be treated similarly but the resulting expressions are much more complicated.

Define the $(T + T' + 1)$ -dimensional vectors as follows:

$$\boldsymbol{\mu}_3 = \begin{bmatrix} 0 \\ b \\ 2b \\ \vdots \\ Tb \\ b' + Tb \\ 2b' + Tb \\ \vdots \\ T'b' + Tb \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ x_1 + a \\ x_2 + 2a + b \\ \vdots \\ x_T + Ta + \frac{(T-1)T}{2}b \\ x_{T+1} + (T+1)a + bT + \frac{(T-1)T}{2}b \\ x_{T+2} + (T+2)a + 2bT + b' + \frac{(T-1)T}{2}b \\ \vdots \\ x_{T+T'} + (T+T')a + bTT' + \frac{(T'-1)T'}{2}b' + \frac{(T-1)T}{2}b \end{bmatrix}, \quad (4.3.2)$$

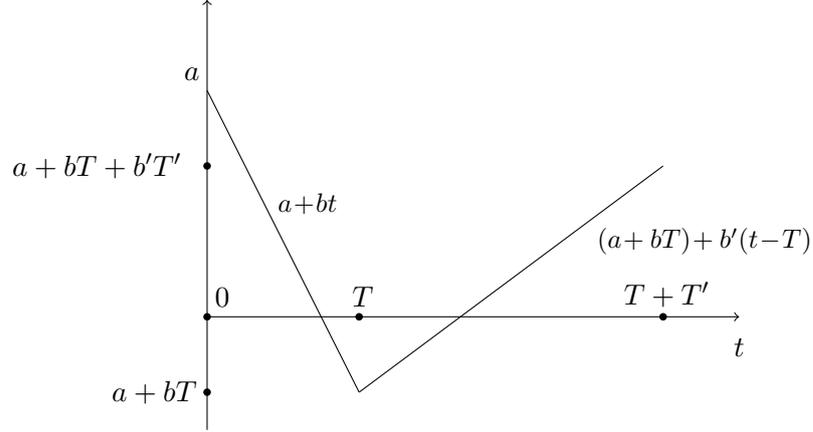


Figure 4.1: Graphical depiction of a general boundary $B_{T,T'}(t; a, b, b')$ with $b < 0$ and $b' > 0$.

$$\mathbf{c}_3 = \begin{bmatrix} x_1 \\ x_2 + a + b \\ x_3 + 2a + 3b \\ \vdots \\ x_T + (T-1)(a+b) + \frac{(T-2)(T-1)}{2}b \\ x_{T+1} + T(a+b) + \frac{(T-1)T}{2}b \\ x_{T+2} + a(T+1) + bT + \frac{(T-1)T}{2}b + b' + Tb \\ \vdots \\ x_{T+T'+1} + a(T+T') + bTT' + \frac{(T'-1)T'}{2}b' + \frac{(T-1)T}{2}b + T'b' + Tb. \end{bmatrix}, \quad (4.3.3)$$

and let a_{3i} and c_{3i} be i -th components of vectors \mathbf{a}_3 and \mathbf{c}_3 respectively ($i = 0, 1, \dots, T+T'$).

Theorem 4.3.1 For $x < a$ and any positive integers T and T' , we have

$$F_{a,b,b'}(T, T' | x) = \frac{1}{\varphi(x)} \int_{-x-a-b}^{\infty} \int_{x_2-a-2b}^{\infty} \cdots \int_{x_T-a-bT}^{\infty} \int_{x_{T+1}-a-bT-b'}^{\infty} \cdots \int_{x_{T+T'+1}-a-bT-b'T'}^{\infty} \exp(-\|\boldsymbol{\mu}_3\|^2/2 + \boldsymbol{\mu}'_3(\mathbf{c}_3 - \mathbf{a}_3)) \det [\varphi(a_{3i} - c_{3j})]_{i,j=0}^{T+T'} dx_{T+T'+1} \dots dx_2. \quad (4.3.4)$$

The proof of Theorem 4.3.1 is included in the appendix, see Section 4.6.2. Note that if $b = b'$ then (4.3.4) reduces to (4.2.5) with $n = T + T'$.

4.3.2 Two particular cases of Theorem 4.3.1

Below we consider two particular cases of Theorem 4.3.1; first, the barrier is $B_{1,1}(t; a, -b, b)$ with $b > 0$; second, the barrier is $B_{1,1}(t; a, 0, -b')$ with $b' > 0$. See Figures 4.2 and 4.3 for a depiction of both barriers. As we demonstrate in Section 4.5, these cases are important for problems of change-point detection.

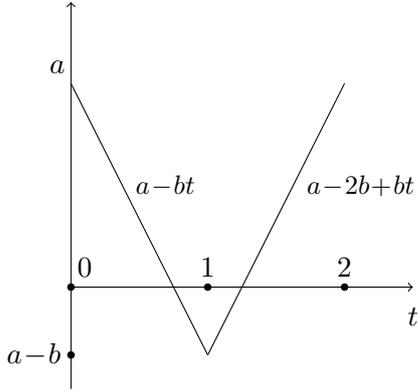


Figure 4.2: Barrier $B_{1,1}(t; a, -b, b)$ with $b > 0$.

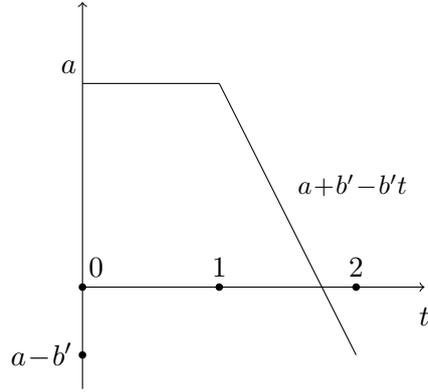


Figure 4.3: Barrier $B_{1,1}(t; a, 0, -b')$ with $b' > 0$.

For the barrier $B_{1,1}(t; a, -b, b)$, an application of Theorem 4.3.1 yields

$$F_{a,-b,b}(1, 1 | x) = \frac{e^{b^2/2-bx}}{\varphi(x)} \int_{-x-a+b}^{\infty} e^{-bx_2} \quad (4.3.5)$$

$$\times \det \begin{bmatrix} \varphi(x) & \varphi(-x_2-a+b) & \Phi(-x_2-a+b) \\ \varphi(a) & \varphi(-x-x_2+b) & \Phi(-x-x_2+b) \\ \varphi(x_2+2a-b+x) & \varphi(a) & \Phi(a) \end{bmatrix} dx_2.$$

For $B_{1,1}(t; a, 0, -b')$, Theorem 4.3.1 provides:

$$F_{a,0,-b'}(1, 1 | x) = \frac{e^{b'^2/2}}{\varphi(x)} \int_{-x-a}^{\infty} \int_{x_2-a+b'}^{\infty} e^{-b'(x_3-x_2)} \quad (4.3.6)$$

$$\times \det \begin{bmatrix} \varphi(x) & \varphi(-x_2-a) & \varphi(-x_3-2a+b') \\ \varphi(a) & \varphi(-x-x_2) & \varphi(-x-x_3-a+b') \\ \varphi(x_2+2a+x) & \varphi(a) & \varphi(x_2-x_3+b') \end{bmatrix} dx_3 dx_2.$$

4.4 Piecewise linear barrier with two changes in slope

4.4.1 Boundary crossing probability

Theorem 4.3.1 can be generalized to the case when we have more than one change in slope. In the general case, the formulas for the first-passage probability become very complicated; they are already rather heavy in the case of one change in slope.

In this section, we consider just one particular barrier with two changes in slope. For real a, b, b', b'' , define the barrier $B(t; a, b, b', b'')$ as

$$B(t; a, b, b', b'') = \begin{cases} a + bt, & t \in [0, 1], \\ a + b + b'(t-1), & t \in [1, 2], \\ a + b + b' + b''(t-2), & t \in [2, 3]. \end{cases}$$

As will be explained in Section 4.5, the corresponding first-passage probability

$$F_{a,b,b',b''}(3 | x) := \Pr(S(t) < B(t; a, b, b', b'') \text{ for all } t \in [0, 3] \mid S(0) = x) \quad (4.4.1)$$

is important for some change-point detection problems.

Define the four-dimensional vectors as follows:

$$\boldsymbol{\mu}_4 = \begin{bmatrix} 0 \\ b \\ b + b' \\ b + b' + b'' \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 0 \\ x_1 + a \\ x_2 + 2a + b \\ x_3 + 3a + 2b + b' \end{bmatrix}, \quad \mathbf{c}_4 = \begin{bmatrix} x_1 \\ x_2 + a + b \\ x_3 + 2a + 2b + b' \\ x_4 + 3a + 3b + 2b' + b'' \end{bmatrix} \quad (4.4.2)$$

and let a_{4i} and c_{4i} be i -th components of vectors \mathbf{a}_4 and \mathbf{c}_4 respectively ($i = 0, 1, 2, 3$).

Theorem 4.4.1 For any real a, b, b', b'' and $x < a$

$$F_{a,b,b',b''}(3|x) = \frac{1}{\varphi(x)} \int_{-x-a-b}^{\infty} \int_{x_2-a-b-b'}^{\infty} \int_{x_3-a-b-b'-b''}^{\infty} \exp(-\|\boldsymbol{\mu}_4\|^2/2 + \boldsymbol{\mu}'_4(\mathbf{c}_4 - \mathbf{a}_4)) \det [\varphi(a_{4i} - c_{4j})]_{i,j=0}^3 dx_4 dx_3 dx_2. \quad (4.4.3)$$

For the proof of Theorem 4.4.1, see Section 4.6.3 in the appendix.

4.4.2 A particular case of Theorem 4.4.1

In this section, we consider a special barrier $B(t; h, 0, -\mu, \mu)$ (depicted in Figure 4.4), which will be used in Section 4.5. In the notation of Theorem 4.4.1, $a = h$, $b = 0$, $b' = -\mu$, $b'' = \mu$ and we obtain

$$F_{h,0,-\mu,\mu}(3|x) = \frac{e^{\mu^2/2}}{\varphi(x)} \int_{-x-h}^{\infty} \int_{x_2-h+\mu}^{\infty} e^{-\mu(x_3-x_2)} dx_3 dx_2 \times \quad (4.4.4)$$

$$\det \begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h+\mu) & \Phi(-x_3-2h+\mu) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h+\mu) & \Phi(-x-x_3-h+\mu) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3+\mu) & \Phi(x_2-x_3+\mu) \\ \varphi(x_3+3h-\mu+x) & \varphi(x_3+2h-\mu-x_2) & \varphi(h) & \Phi(h) \end{bmatrix}.$$

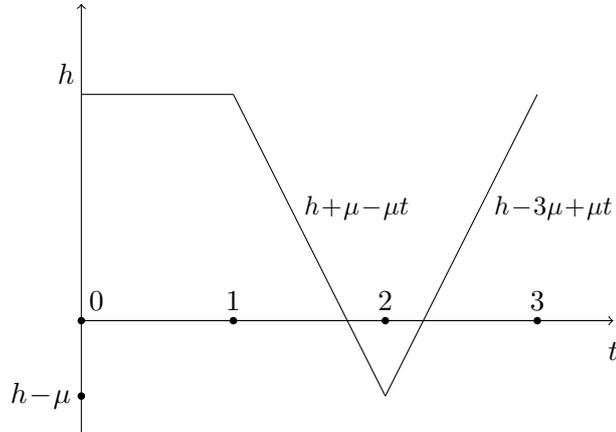


Figure 4.4: Barrier $B(t; h, 0, -\mu, \mu)$ with $\mu > 0$.

4.4.3 Another linear barrier with two changes in slope

For real h and μ , define the barrier $B(t; h, 0, 0, -\mu, \mu)$ by

$$B(t; h, 0, 0, -\mu, \mu) = \begin{cases} h, & t \in [0, 2], \\ h - \mu(t - 2), & t \in [2, 3], \\ h - \mu + \mu(t - 3), & t \in [3, 4]. \end{cases}$$

The barrier $B(t; h, 0, 0, -\mu, \mu)$ looks similar to the barrier depicted in Figure 4.4, except the constant part is two units long. The corresponding first-passage probability

$$F_{h,0,0,-\mu,\mu}(4|x) := \Pr(S(t) < B(t; h, 0, 0, -\mu, \mu) \text{ for all } t \in [0, 4] \mid S(0) = x) \quad (4.4.5)$$

will be important in Section 4.5.

Theorem 4.4.2 *For any real h , μ and $x < h$*

$$F_{h,0,0,-\mu,\mu}(4|x) = \frac{e^{\mu^2/2}}{\varphi(x)} \int_{-x-h}^{\infty} dx_2 \int_{x_2-h}^{\infty} dx_3 \int_{x_3-h+\mu}^{\infty} dx_4 e^{-\mu(x_4-x_3)} \times \quad (4.4.6)$$

$$\det \begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h) & \varphi(-x_4-3h+\mu) & \Phi(-x_4-3h+\mu) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h) & \varphi(-x-x_4-2h+\mu) & \Phi(-x-x_4-2h+\mu) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3) & \varphi(x_2-x_4-h+\mu) & \Phi(x_2-x_4-h+\mu) \\ \varphi(x_3+3h+x) & \varphi(x_3+2h-x_2) & \varphi(h) & \varphi(x_3+\mu-x_4) & \Phi(x_3+\mu-x_4) \\ \varphi(x_4+4h-\mu+x) & \varphi(x_4+3h-\mu-x_2) & \varphi(x_4+2h-\mu-x_3) & \varphi(h) & \Phi(h) \end{bmatrix}.$$

The proof of Theorem 4.4.2 is very similar to the proof of Theorem 4.4.1.

4.5 Application to change-point detection

4.5.1 Formulation of the problem

In this section, we illustrate the natural appearance of the first-passage probabilities for the Slepian process $S(t)$ for piece-wise linear barriers and in particular the barriers considered in Sections 4.3.2 and 4.4.2.

Suppose one can observe the stochastic process $X(t)$ ($t \geq 0$) governed by the stochastic differential equation

$$dX(t) = \mu \mathbb{1}_{\{\nu \leq t < \nu+l\}} dt + dW(t), \quad (4.5.1)$$

where $\nu > 0$ is the unknown (non-random) change-point and $\mu \neq 0$ is the drift magnitude during the ‘epidemic’ period of duration l with $0 < l < \infty$; μ and l may be known or unknown. The classical change-point detection problem of finding a change in drift of a Wiener process is the problem (4.5.1) with $l = \infty$; that is, when the change (if occurred) is permanent, see for example [76, 94, 98, 99].

In (4.5.1), under the null hypothesis \mathbb{H}_0 , we assume $\nu = \infty$ meaning that the process $dX(t)$ has zero mean for all $t \geq 0$. On the other hand, under the alternative hypothesis \mathbb{H}_1 , $\nu < \infty$. In the definition of the test power, we will assume that ν is

large. However, for the tests discussed below to be well-defined and approximations to be accurate, we only need $\nu \geq 1$ (under \mathbb{H}_1).

In this section, we only consider the case of known l , in which case we can assume $l = 1$ (otherwise we change the time-scale by $t \rightarrow t/l$ and the barrier by $B \rightarrow B/\sqrt{l}$). When testing for an epidemic change on a fixed interval $[0, T]$ with l unknown, one possible approach is to construct the test statistic on the base of $\max_{0 < s < t < T} [W(t) - W(s)]$, the maximum over all possible choices of l and locations. This idea was discussed in [118], where asymptotic approximations are offered. The case when l is unknown is more complicated and the first-passage probabilities that have to be used are more involved.

We define the test statistic used to monitor the epidemic alternative as

$$S_1(t) = \int_t^{t+1} dX(t) \quad t \geq 0.$$

The stopping rule for $S_1(t)$ is defined as follows

$$\tau(h) = \inf\{t : S_1(t) \geq h\}, \quad (4.5.2)$$

where the threshold h is chosen to satisfy the average run length (ARL) constraint $\mathbb{E}_0(\tau(h)) = C$ for some (usually large) fixed C (here \mathbb{E}_0 denote the expectation under the null hypothesis). Since l is known, for any $\mu > 0$ the test with the stopping rule (4.5.2) is optimal in the sense of the Abstract Neyman-Pearson lemma, see Theorem 2, [38, p 110].

The process $S_1(t) - \mathbb{E}S_1(t) = W(t+1) - W(t)$ is stochastically equivalent to the Slepian process $S(t)$ of (4.1.1). Under \mathbb{H}_0 , $\mathbb{E}S_1(t) = 0$ for all $t \geq 0$ and under \mathbb{H}_1 we have

$$\mathbb{E}S_1(t) = \begin{cases} \mu(t - \nu + 1) & \text{for } \nu - 1 < t \leq \nu \\ \mu(1 - t + \nu) & \text{for } \nu < t \leq \nu + 1 \\ 0 & \text{otherwise.} \end{cases}$$

4.5.2 Approximation for $\mathbb{E}_0(\tau(h))$

To construct accurate approximations for $\mathbb{E}_0(\tau(h))$, we shall utilise some of the approximations derived in Section 3.2.5 of Chapter 3. The resulting approximation will have the same form as the approximation provided in Section 2.8 of Chapter 2, without the discrete time correction. Consider the unconditional probability (taken with respect to the standard normal distribution):

$$F_{h,0}(T) := \int_{-\infty}^h F_{h,0}(T|x)\varphi(x)dx.$$

Under \mathbb{H}_0 , the distribution of $\tau(h)$ has the form $(1 - \Phi(h))\delta_0(ds) + q_h(s)ds$, $s \geq 0$, where $\delta_0(ds)$ is the delta-measure concentrated at 0 and

$$q_h(s) = -\frac{d}{ds}F_{h,0}(s), \quad 0 < s < \infty$$

is the first-passage density. This yields

$$\mathbb{E}_0(\tau(h)) = \int_0^\infty sq_h(s)ds. \quad (4.5.3)$$

There is no easy computationally convenient formula for $q_h(t)$ as expressions for $F_{h,0}(s)$ are very complex. One of the simplest (yet very accurate) approximation for $F_{h,0}(s)$ takes the form:

$$F_{h,0}(T) \simeq F_{h,0}(2) \cdot \lambda(h)^{T-2}, \quad \text{for all } T > 0, \quad (4.5.4)$$

with $\lambda(h) = F_{h,0}(2)/F_{h,0}(1)$. This corresponds to Approximation 4 in Chapter 3. Using (4.5.4), we approximate the density $q_h(s)$ by

$$q_h(s) \simeq -F_{h,0}(2) \log[\lambda(h)] \cdot \lambda(h)^{s-2}, \quad 0 < s < \infty.$$

Subsequent evaluation of the integral in (4.5.3) yields the approximation

$$\mathbb{E}_0(\tau(h)) \simeq -\frac{F_{h,0}(2)}{\lambda(h)^2 \log[\lambda(h)]}. \quad (4.5.5)$$

Numerical study shows that the approximation (4.5.5) is very accurate for all $h \geq 3$. Setting $h = 3.63$ in (4.5.5) results in $C \simeq 500$.

4.5.3 Approximating the power of the test

In this section we formulate several approximations for the power of the test (4.5.2) which can be defined as

$$\mathcal{P}(h, \mu) := \lim_{\nu \rightarrow \infty} \mathbb{P}_1 \{S_1(t) \geq h \text{ for at least one } t \in [\nu - 1, \nu + 1] \mid \tau(h) > \nu - 1\}, \quad (4.5.6)$$

where \mathbb{P}_1 denotes the probability measure under the alternative hypothesis. Define the piecewise linear barrier $Q_\nu(t; h, \mu)$ as follows

$$Q_\nu(t; h, \mu) = h - \mu \max\{0, 1 - |t - \nu|\}.$$

The barrier $Q_\nu(t; h, \mu)$ is visually depicted in Figure 4.5. The power of the test with the stopping rule (4.5.2) is then

$$\mathcal{P}(h, \mu) = \lim_{\nu \rightarrow \infty} \mathbb{P} \{S(t) \geq Q_\nu(t; h, \mu) \text{ for at least one } t \in [\nu - 1, \nu + 1] \mid \tau(h) > \nu - 1\}.$$

Consider the barrier $B(t; h, 0, -\mu, \mu)$ of Section 4.4 with $t \in [0, 3]$. Define the conditional first-passage probability

$$\begin{aligned} \gamma_3(x, h, \mu) &:= \mathbb{P}\{S(t) \geq B(t; h, 0, -\mu, \mu) \text{ for some } t \in [1, 3] \mid S(0) = x; S(t) < h, \forall t \in [0, 1]\} \\ &= 1 - \frac{\mathbb{P}\{S(t) < B(t; h, 0, -\mu, \mu) \text{ for all } t \in [0, 3] \mid S(0) = x\}}{\mathbb{P}\{S(t) < h \text{ for all } t \in [0, 1] \mid S(0) = x\}} = 1 - \frac{F_{h,0,-\mu,\mu}(3|x)}{F_{h,0}(1|x)}. \end{aligned} \quad (4.5.7)$$

The denominator in (4.5.7) is very simple to compute, see (4.2.6) with $b = 0$ and $a = h$. The numerator in (4.5.7) can be computed by (4.4.4). Computation of $\gamma_3(x, h, \mu)$ requires numerical evaluation of a two-dimensional integral, which is not difficult. Recall that in this thesis we use the convention that $\Phi(x)$ is explicit

and not an integral. This is because $\Phi(x)$ can be easily evaluated by all statistical software.

The first approximation to the power $\mathcal{P}(h, \mu)$ is $\gamma_3(0, h, \mu)$. In view of (4.1.1) the process $S(t)$ forgets the past after one unit of time hence quickly reaches the stationary behaviour under the condition $S(t) < h$ for all $t < \nu - 1$. By approximating $\mathcal{P}(h, \mu)$ with $\gamma_3(0, h, \mu)$, we assume that one unit of time is almost enough for $S(t)$ to reach this stationary state. In Figure 4.6, we plot the ratio $\gamma_3(x, h, \mu)/\gamma_3(0, h, \mu)$ as a function of x for $h = 3$ and $\mu = 3$. Since the ratio is very close to 1 for all considered x , this verifies that the probability $\gamma_3(x, h, \mu)$ changes very little as x varies implying that the values of $S(t)$ at $t = \nu - 2$ have almost no effect on the probability $\gamma_3(x, h, \mu)$. This allows us to claim that the accuracy $|\mathcal{P}(h, \mu) - \gamma_3(0, h, \mu)|$ of the approximation $\mathcal{P}(h, \mu) \simeq \gamma_3(0, h, \mu)$ is smaller than 10^{-4} for all $h \geq 3$.

Consider the barrier $B(t; h, 0, 0, -\mu, \mu)$ of Section 4.4.3 with $t \in [0, 4]$. Define the conditional first-passage probability

$$\begin{aligned} \gamma_4(x, h, \mu) &:= \mathbb{P}\{S(t) \geq B(t; h, 0, 0, -\mu, \mu) \text{ for some } t \in [2, 4] \mid S(0) = x, S(t) < h, \forall t \in [0, 2]\} \\ &= 1 - \frac{\mathbb{P}\{S(t) < B(t; h, 0, 0, -\mu, \mu) \text{ for all } t \in [0, 4] \mid S(0) = x\}}{\mathbb{P}\{S(t) < h \text{ for all } t \in [0, 2] \mid S(0) = x\}} = 1 - \frac{F_{h,0,0,-\mu,\mu}(4|x)}{F_{h,0}(2|x)}. \end{aligned}$$

The numerator in $\gamma_4(x, h, \mu)$ requires numerical evaluation of the three-dimensional integral in (4.4.6). The denominator can be computed using Theorem 4.2.1 with $a = h$ and $b = 0$. The second approximation to the power $\mathcal{P}(h, \mu)$ is $\gamma_4(0, h, \mu)$. The accuracy of the approximation $\mathcal{P}(h, \mu) \simeq \gamma_4(0, h, \mu)$ is smaller than 10^{-6} for all $h \geq 3$ and $\mu \geq 0$. In particular, $|\gamma_4(1, 3, 3)/\gamma_4(-1, 3, 3) - 1| < 10^{-7}$, compare this with Figure 4.6. For $h = 3.11$ and hence $C \simeq 100$, we have $|\gamma_4(0, h, 3)/\gamma_3(0, h, 3) - 1| < 3 \cdot 10^{-5}$ and $|\gamma_4(0, h, 4)/\gamma_3(0, h, 4) - 1| < 6 \cdot 10^{-6}$. The approximation $\gamma_3(0, h, \mu)$ is the most favourable since it is almost as precise as $\gamma_4(0, h, \mu)$ but computationally $\gamma_3(0, h, \mu)$ is much cheaper.

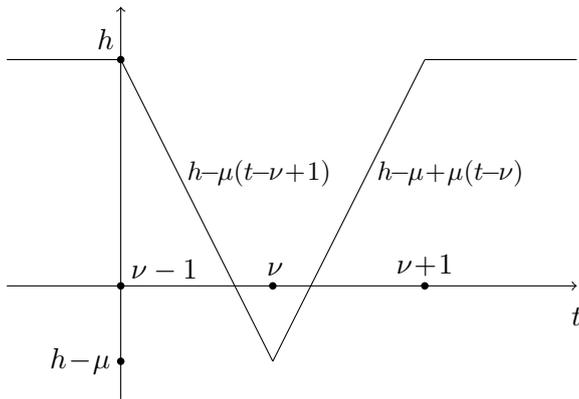


Figure 4.5: Graphical depiction of the boundary $Q_\nu(t; h, \mu)$.

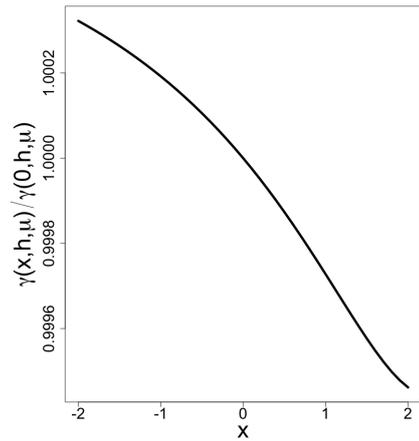


Figure 4.6: Ratio $\gamma(x, h, \mu)/\gamma(0, h, \mu)$ for $h = 3$ and $\mu = 3$.

As seen from Figures 4.2 and 4.4, the barrier $B_{1,1}(t; h, -\mu, \mu)$ is the main component of the barrier $B(t; h, 0, -\mu, \mu)$. Instead of using the approximation $\mathcal{P}(h, \mu) \simeq$

$\gamma_3(0, h, \mu)$ it is therefore tempting to use a simpler approximation $\mathcal{P}(h, \mu) \simeq \gamma_2(0, h, \mu)$, where

$$\gamma_2(x, h, \mu) := \mathbb{P}\{S(t) \geq B_{1,1}(t; h, -\mu, \mu) \text{ for some } t \in [0, 2] \mid S(0) = x\} = 1 - F_{h, -\mu, \mu}(1, 1 \mid x).$$

To compute values of $\gamma_2(0, h, \mu)$ we only need to evaluate the one-dimensional integral in (4.3.5) with $b = \mu$.

To assess the impact of the final line-segment in the barrier $B(t; h, 0, -\mu, \mu)$ on the power (the line-segment with gradient μ in Fig 4.5, $t \in [\nu, \nu + 1]$), let

$$\begin{aligned} \gamma_1(x, h, \mu) &:= \mathbb{P}\{S(t) \geq B_{1,1}(t; h, 0, -\mu) \text{ for some } t \in [1, 2] \mid S(0) = x, S(t) < h, \forall t \in [0, 1]\} \\ &= 1 - \frac{\mathbb{P}\{S(t) < B_{1,1}(t; h, 0, -\mu) \text{ for all } t \in [0, 2] \mid S(0) = x\}}{\mathbb{P}\{S(t) < h \text{ for all } t \in [0, 1] \mid S(0) = x\}} = 1 - \frac{F_{h, 0, -\mu}(1, 1 \mid x)}{F_{h, 0}(1 \mid x)}. \end{aligned}$$

Then we make the approximation $\mathcal{P}(h, \mu) \simeq \gamma_1(0, h, \mu)$, where the quantity $F_{h, 0, -\mu}(1, 1 \mid 0)$ can be computed using (4.3.6) with $b' = \mu$. The denominator can be computed using (4.2.6) with $b = 0$ and $a = h$.

In Table 4.1, we provide values of $\mathcal{P}(h, \mu)$, $\gamma_2(0, h, \mu)$ and $\gamma_1(0, h, \mu)$ for different μ , where the values of h have been chosen to satisfy $\mathbb{E}_0(\tau(h)) = C$ for $C = 100, 500, 1000$; see (4.5.5) regarding computation of the ARL $\mathbb{E}_0(\tau(h))$. Since the values in Table 4.1 are given to three decimal places, these values of $\mathcal{P}(h, \mu)$ can be obtained from either $\gamma_3(0, h, \mu)$ or $\gamma_4(0, h, \mu)$; both of these two approximations provide a better accuracy than 3 decimal places. Comparing the entries of Table 4.1 we can observe that the quality of the approximation $\mathcal{P}(h, \mu) \simeq \gamma_2(0, h, \mu)$ is rather good, especially for large μ . By looking at the columns corresponding to $\gamma_1(0, h, \mu)$, one can also see the expected diminishing impact which the final line-segment in $B(t; h, 0, -\mu, \mu)$ has on power, as μ increases. However, for small μ the contribution of this part of the barrier to power is significant suggesting it is not be sensible to approximate the power of the test considered here with $\gamma_1(0, h, \mu)$.

μ	$h = 3.11, C \simeq 100$			$h = 3.63, C \simeq 500$			$h = 3.83, C \simeq 1000$		
	\mathcal{P}	γ_2	γ_1	\mathcal{P}	γ_2	γ_1	\mathcal{P}	γ_2	γ_1
2	0.305	0.292	0.239	0.138	0.131	0.104	0.096	0.090	0.071
2.25	0.388	0.375	0.315	0.195	0.187	0.152	0.140	0.134	0.108
2.5	0.476	0.464	0.402	0.264	0.255	0.213	0.198	0.191	0.157
2.75	0.568	0.557	0.494	0.345	0.336	0.288	0.269	0.262	0.221
3	0.656	0.647	0.587	0.434	0.426	0.373	0.351	0.344	0.297
3.25	0.737	0.730	0.676	0.527	0.520	0.466	0.442	0.435	0.385
3.5	0.808	0.802	0.757	0.620	0.613	0.561	0.536	0.530	0.479
3.75	0.865	0.861	0.825	0.706	0.701	0.653	0.629	0.623	0.574
4	0.910	0.907	0.880	0.782	0.778	0.737	0.715	0.710	0.666
4.25	0.943	0.941	0.922	0.846	0.843	0.810	0.790	0.787	0.749
4.5	0.965	0.964	0.951	0.896	0.894	0.869	0.852	0.850	0.819
4.75	0.980	0.980	0.971	0.933	0.932	0.913	0.901	0.899	0.876
5	0.989	0.989	0.984	0.959	0.958	0.946	0.937	0.936	0.919

Table 4.1: $\mathcal{P}(h, \mu)$, $\gamma_2(0, h, \mu)$ and $\gamma_1(0, h, \mu)$ for different μ for three choices of ARL.

To summarize the results of this section, for approximating the power function $\mathcal{P}(h, \mu)$, we propose one the following two approximations: a very accurate approximation $\gamma_3(0, h, \mu)$ requiring numerical evaluation of a two-dimensional integral and $\gamma_2(0, h, \mu)$, a less accurate but simpler approximation requiring evaluation of a one-dimensional integral only. The approximation $\mathcal{P}(h, \mu) \simeq \gamma_4(0, h, \mu)$ is extremely accurate but too costly whereas the approximation $\gamma_1(0, h, \mu)$ is less accurate than $\gamma_2(0, h, \mu)$ but slightly cheaper, requiring the numerical evaluation of a two-dimensional integral. The approximation $\mathcal{P}(h, \mu) \simeq \gamma_1(0, h, \mu)$ has been studied mainly for assessing the impact which the final line-segment in $B(t; h, 0, -\mu, \mu)$ has on the power.

4.6 Appendix

4.6.1 Proof of Theorem 4.2.2

Using (4.1.1), the first-passage probability $F_{a,b}(T | x)$ can be equivalently expressed as follows

$$\begin{aligned}
& F_{a,b}(T | x) \\
&= \Pr\{W(t) - W(t+1) < a + bt \text{ for all } t \in [0, m + \theta] \mid W(0) - W(1) = x\} \\
&= \Pr(W(t) - W(t+1) < a + bt, W(t+1) - W(t+2) < a + b(t+1), \dots, \\
&\quad W(t+m) - W(t+m+1) < a + b(t+m) \text{ for all } t \in [0, \theta] \text{ and} \\
&\quad W(\tau + \theta) - W(\tau + \theta + 1) < a + b\theta + b\tau, W(\tau + \theta + 1) - W(\tau + \theta + 2) \\
&\quad < a + b + b\theta + b\tau, \dots, W(\tau + (m-1) + \theta) - W(\tau + m + \theta) < \\
&\quad a + b\theta + (m-1)b + b\tau \text{ for all } \tau \in [0, 1 - \theta] \mid W(0) - W(1) = x) \\
&= \Pr\left\{W(t) < W(t+1) + a + bt < \dots < W(t+m+1) + (m+1)(a+bt) + \frac{(m+1)m}{2}b \right. \\
&\quad \left. \forall t \in [0, \theta] \text{ and } W(\tau + \theta) < W(\tau + \theta + 1) + a + b\theta + b\tau < \dots < \right. \\
&\quad \left. W(\tau + \theta + m) + m(a + b\theta + b\tau) + \frac{(m-1)m}{2}b \forall \tau \in [0, 1 - \theta] \mid W(0) - W(1) = x\right\}.
\end{aligned}$$

Let Ω be the event

$$\begin{aligned}
\Omega = & \left\{ W(t) < W(t+1) + a + bt < \dots < W(t+m+1) + (m+1)(a+bt) + \frac{(m+1)m}{2}b \right. \\
& \left. \forall t \in [0, \theta] \text{ and } W(\tau + \theta) < W(\tau + \theta + 1) + a + b\theta + b\tau < \dots < W(\tau + \theta + m) + \right. \\
& \left. m(a + b\theta + b\tau) + \frac{(m-1)m}{2}b \forall \tau \in [0, 1 - \theta] \right\}.
\end{aligned}$$

By integrating out over the values u_i and v_i of W at times i and $i+\theta$, $i = 0, 1, \dots, m+1$, by the law of total probability we have

$$\begin{aligned}
F_{a,b}(T|x) &= \int \dots \int \Pr\{\Omega \mid W(0)=u_0, \dots, W(m+1)=u_{m+1}, W(\theta)=v_0, \dots, \\
&\quad W(m+1+\theta)=v_{m+1}, W(0)-W(1)=x\} \times \\
&\quad \Pr\{W(0) \in du_0, \dots, W(m+1) \in du_{m+1}, W(\theta) \in dv_0, \dots, \\
&\quad W(m+1+\theta) \in dv_{m+1} \mid W(0)-W(1)=x\}. \quad (4.6.1)
\end{aligned}$$

Since $W(0) - W(1) = x$ and $W(0) = 0$, we have $W(1) = x_1 = -x$. Define the processes

$$\begin{aligned}
W_i(t) &= W(t+i) + i(a+bt) + \frac{(i-1)i}{2}b, \quad 0 \leq t \leq \theta, \quad i = 0, 1, \dots, m+1, \\
W'_j(\tau) &= W(\tau+\theta+j) + j(a+b\theta+b\tau) + \frac{(j-1)j}{2}b, \quad 0 \leq \tau \leq 1-\theta, \quad j = 0, 1, \dots, m.
\end{aligned}$$

Then the event Ω can be equivalently expressed as $\Omega = \Omega_1 \cap \Omega_2$ with

$$\begin{aligned}
\Omega_1 &= \{W_0(t) < W_1(t) < \dots < W_{m+1}(t) \text{ for all } t \in [0, \theta]\}, \\
\Omega_2 &= \{W'_0(\tau) < W'_1(\tau) < \dots < W'_m(\tau) \text{ for all } \tau \in [0, 1-\theta]\}.
\end{aligned}$$

Under the conditioning introduced in (4.6.1) we have for $i = 0, 1, \dots, m+1$ and $j = 0, 1, \dots, m$:

$$\begin{aligned}
W_i(0) &= W(i) + ia + \frac{(i-1)i}{2}b = u_i + ia + \frac{(i-1)i}{2}b, \\
W_i(\theta) &= W(i+\theta) + i(a+b\theta) + \frac{(i-1)i}{2}b = v_i + i(a+b\theta) + \frac{(i-1)i}{2}b, \\
W'_j(0) &= W(j+\theta) + j(a+b\theta) + \frac{(j-1)j}{2}b = v_j + j(a+b\theta) + \frac{(j-1)j}{2}b, \\
W'_j(1-\theta) &= W(j+1) + j(a+b) + \frac{(j-1)j}{2}b = u_{j+1} + j(a+b) + \frac{(j-1)j}{2}b.
\end{aligned}$$

Now under the above conditioning, the processes are independent and so the conditional probability of Ω in (4.6.1) becomes a product of the conditional probabilities of Ω_1 and Ω_2 . Therefore, (4.6.1) becomes

$$\begin{aligned}
F_{a,b}(T|x) &= \\
&\int \dots \int \Pr\left\{\Omega_1 \mid W_i(0) = u_i + ia + \frac{(i-1)i}{2}b, W_i(\theta) = v_i + i(a+b\theta) + \frac{(i-1)i}{2}b, \right. \\
&\quad \left. (0 \leq i \leq m+1)\right\} \times \Pr\left\{\Omega_2 \mid W'_j(0) = v_j + j(a+b\theta) + \frac{(j-1)j}{2}b \right. \\
&\quad \left. W'_j(1-\theta) = u_{j+1} + j(a+b) + \frac{(j-1)j}{2}b \quad (0 \leq j \leq m)\right\} \times \Pr\{W(0) \in du_0 \\
&\quad \dots, W(m+1) \in du_{m+1}, W(\theta) \in dv_0, \dots, W(m+1+\theta) \in dv_{m+1} \\
&\quad \mid W(0)-W(1)=x\}. \quad (4.6.2)
\end{aligned}$$

The region of integration for the variables u_i in (4.6.2) is determined from the following chain of inequalities:

$$-x - a < u_2 + 2a + b < \dots < u_m + ma + \frac{(m-1)m}{2}b < u_{m+1} + (m+1)a + \frac{(m+1)m}{2}b.$$

Whence, the upper limit of integration with respect to u_{i+1} is infinity and the lower limit for the integral with respect to u_{i+1} , $i = 1, \dots, m$ is given by the formula $u_i - a - ib$. For the variables v_j in (4.6.2), we have the following chain of inequalities

$$v_0 < v_1 + a + b\theta < \dots < v_m + m(a + b\theta) + \frac{(m-1)m}{2}b < v_{m+1} + (m+1)(a + b\theta) + \frac{(m+1)m}{2}b.$$

Once again, the upper limit of integration with respect to v_{i+1} is infinity and the lower limit for the integral with respect to v_{i+1} ($i = 0, \dots, m$) is $v_i - a - b\theta - ib$. For v_0 , the upper and lower limits of integration are infinite. Now using (4.2.3) with $n = m + 1$ we obtain

$$\begin{aligned} \Pr \left\{ \Omega_1 \mid W_i(0) = u_i + ia + \frac{(i-1)i}{2}b, W_i(\theta) = v_i + i(a + b\theta) + \frac{(i-1)i}{2}b, (0 \leq i \leq m+1) \right\} \\ = \exp(-\theta \|\boldsymbol{\mu}_1\|^2/2 + \boldsymbol{\mu}'_1(\mathbf{c}_1 - \mathbf{a}_1)) \det[\varphi_\theta(a_{1i} - c_{1j})]_{i,j=0}^{m+1} / \prod_{i=0}^{m+1} \varphi_\theta(a_{1i} - c_{1i} + \theta\mu_{1i}), \end{aligned}$$

where $\varphi_\theta(\cdot)$ is given in (4.2.1), \mathbf{a}_1 and \mathbf{c}_1 are given in (4.2.7). Similarly, using (4.2.3) with $n = m$ we have

$$\begin{aligned} \Pr \left\{ \Omega_2 \mid W'_j(0) = v_j + j(a + b\theta) + \frac{(j-1)j}{2}b, W'_j(1-\theta) = u_{j+1} + j(a + b) + \frac{(j-1)j}{2}b, (0 \leq j \leq m) \right\} \\ = \exp(-(1-\theta) \|\boldsymbol{\mu}_2\|^2/2 + \boldsymbol{\mu}'_2(\mathbf{c}_2 - \mathbf{a}_2)) \det[\varphi_{1-\theta}(a_{2i} - c_{2j})]_{i,j=0}^m \\ / \prod_{i=0}^m \varphi_{1-\theta}(a_{2i} - c_{2i} + (1-\theta)\mu_{2i}), \end{aligned}$$

where $\varphi_{1-\theta}(\cdot)$ is given in (4.2.1), \mathbf{a}_2 and \mathbf{c}_2 are given in (4.2.8). The third probability in the right-hand side of (4.6.2) is simply

$$\frac{1}{\varphi(x)} \prod_{j=0}^m \prod_{i=0}^{m+1} \varphi_\theta(u_i - v_i) \varphi_{1-\theta}(v_j - u_{j+1}) dv_j du_{j+1}.$$

By noticing

$$\prod_{j=0}^m \prod_{i=0}^{m+1} \varphi_\theta(a_{1i} - c_{1i} + \theta\mu_{1i}) \varphi_{1-\theta}(a_{2j} - c_{2j} + (1-\theta)\mu_{2j}) = \prod_{j=0}^m \prod_{i=0}^{m+1} \varphi_\theta(u_i - v_i) \varphi_{1-\theta}(v_j - u_{j+1})$$

and collating all terms, we obtain the result. \square

4.6.2 Proof of Theorem 4.3.1 (and Theorem 4.2.1)

We recall that the proof of Theorem 4.2.1 can be obtained by setting $n = T$ and $T' = 0$ in the following proof of Theorem 4.3.1.

Using (4.1.1) we rewrite $F_{a,b,b'}(T, T' | x)$ as

$$\begin{aligned}
& F_{a,b,b'}(T, T' | x) \\
&= \Pr\{W(t) - W(t+1) < a + bt \text{ for all } t \in [0, T], \\
&\quad W(t) - W(t+1) < a + bT + b'(t - T) \text{ for all } t \in [T, T + T'] \mid W(0) - W(1) = x\} \\
&= \Pr\{W(t) - W(t+1) < a + bt, W(t+1) - W(t+2) < a + b(t+1), \dots, \\
&\quad W(t+T-1) - W(t+T) < a + b(t+T-1), W(t+T) - W(t+T+1) < a + bT + b't, \\
&\quad W(t+T+1) - W(t+T+2) < a + bT + b'(t+1) \dots, \\
&\quad W(t+T+T'-1) - W(t+T+T') < a + bT + b'(t+T'-1) \forall t \in [0, 1] \\
&\quad \mid W(0) - W(1) = x\} \\
&= \Pr\left\{W(t) < W(t+1) + a + bt < \dots < W(t+T) + T(a + bt) + \frac{(T-1)T}{2}b \right. \\
&\quad < W(t+T+1) + a(T+1) + bT + \frac{(T-1)T}{2}b + (b' + Tb)t < \dots < \\
&\quad W(t+T+T') + a(T+T') + bTT' + \frac{(T'-1)T'}{2}b' + \frac{(T-1)T}{2}b + (T'b' + Tb)t \\
&\quad \left. \text{for all } t \in [0, 1] \mid W(0) - W(1) = x\right\}.
\end{aligned}$$

Let Ω be the event defined as follows

$$\begin{aligned}
\Omega = & \left\{W(t) < W(t+1) + a + bt < \dots < W(t+T) + T(a + bt) + \frac{(T-1)T}{2}b \right. \\
& < W(t+T+1) + a(T+1) + bT + \frac{(T-1)T}{2}b + (b' + Tb)t < \dots < \\
& W(t+T+T') + a(T+T') + bTT' + \frac{(T'-1)T'}{2}b' + \frac{(T-1)T}{2}b + (T'b' + Tb)t \\
& \left. \text{for all } t \in [0, 1]\right\},
\end{aligned}$$

and let $x_i = W(i)$, $i = 0, \dots, T + T' + 1$. Integrating out over the values x_i , by the law of total probability we obtain:

$$\begin{aligned}
& F_{a,b,b'}(T, T' | x) \\
&= \int \dots \int \Pr\{\Omega \mid W(0) = x_0, \dots, W(T + T' + 1) = x_{T+T'+1}, W(0) - W(1) = x\} \\
&\quad \times \Pr\{W(0) \in dx_0, \dots, W(T + T' + 1) \in dx_{T+T'+1} \mid W(0) - W(1) = x\}. \quad (4.6.3)
\end{aligned}$$

Note that $W(1) = x_1 = -x$, since $W(0) - W(1) = x$ and $W(0) = 0$. Define the following processes which take different forms depending on the value of i :

$$\begin{aligned}
W_i(t) &= W(t+i) + i(a + bt) + \frac{(i-1)i}{2}b, \text{ for } 0 \leq i \leq T; \\
W_i(t) &= W(t+i) + ai + bT(i-T) + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b + \{(i-T)b' + Tb\}t,
\end{aligned}$$

for $T+1 \leq i \leq T+T'$, with $0 \leq t \leq 1$ for all processes. The event Ω can now be expressed as

$$\Omega = \{W_0(t) < W_1(t) < \dots < W_T(t) < \dots < W_{T+T'}(t) \text{ for all } t \in [0, 1]\}. \quad (4.6.4)$$

Under the conditioning introduced in (4.6.3), depending on the size of i we have: for $0 \leq i \leq T$

$$W_i(0) = x_i + ia + \frac{(i-1)i}{2}b, \quad W_i(1) = x_{i+1} + i(a+b) + \frac{(i-1)i}{2}b;$$

and for $T+1 \leq i \leq T+T'$

$$\begin{aligned} W_i(0) &= x_i + ai + bT(i-T) + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b, \\ W_i(1) &= x_{i+1} + ai + bT(i-T) + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b + (i-T)b' + Tb. \end{aligned}$$

Whence (4.6.3) can be expressed as

$$\begin{aligned} &F_{a,b,b'}(T, T' | x) \\ &= \int \dots \int \Pr \left\{ \Omega \mid W_i(0) = x_i + ia + \frac{(i-1)i}{2}b, W_i(1) = x_{i+1} \right. \\ &\quad \left. + i(a+b) + \frac{(i-1)i}{2}b \quad (0 \leq i \leq T), W_i(0) = x_i + ai + bT(i-T) \right. \\ &\quad \left. + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b, W_i(1) = x_{i+1} + ai + bT(i-T) \right. \\ &\quad \left. + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b + (i-T)b' + Tb \right. \\ &\quad \left. (T \leq i \leq T+T'), W_0(0) - W_0(1) = x \right\} \times \\ &\quad \Pr\{W(0) \in dx_0, \dots, W(T+T'+1) \in dx_{T+T'+1} \mid W(0) - W(1) = x\}. \quad (4.6.5) \end{aligned}$$

The region of integration in (4.6.5) is determined from the following inequalities which ensure that the inequalities in (4.6.4) hold at $t=0$ and $t=1$:

$$\begin{aligned} x_1 &< \dots < x_{T+1} + T(a+b) + \frac{(T-1)T}{2}b < x_{T+2} + a(T+1) + bT + \frac{(T-1)T}{2}b + b' + Tb \\ &< \dots < x_{T+T'+1} + a(T+T') + bTT' + \frac{(T'-1)T'}{2}b' + \frac{(T-1)T}{2}b + T'b' + Tb. \end{aligned}$$

From this, the upper limit of integration is infinity for all x_i . For $0 \leq i \leq T+1$, the lower limit for x_i is $x_{i-1} - a - (i-1)b$. For $T+2 \leq i \leq T+T'+1$, the lower limit for x_i is $x_{i-1} - a - bT - b'(i-T-1)$. Since the conditioned Brownian motion

processes $W_i(t)$ are independent, application of (4.2.3) with $n = T + T'$ provides

$$\begin{aligned} \Pr \left\{ \Omega \left| \begin{aligned} W_i(0) &= x_i + ia + \frac{(i-1)i}{2}b, W_i(1) = x_{i+1} + i(a+b) + \frac{(i-1)i}{2}b \quad (0 \leq i \leq T) \\ W_i(0) &= x_i + ai + bT(i-T) + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b, \\ W_i(1) &= x_{i+1} + ai + bT(i-T) + \frac{(i-T-1)(i-T)}{2}b' + \frac{(T-1)T}{2}b + (i-T)b' + Tb \\ (T \leq i \leq T+T'), W_0(0) - W_0(1) &= x \end{aligned} \right. \right\} \\ &= \exp(-\|\boldsymbol{\mu}_3\|^2/2 + \boldsymbol{\mu}'_3(\mathbf{c}_3 - \mathbf{a}_3)) \det[\varphi(a_{3i}, c_{3j})]_{i,j=0}^{T'+T} / \prod_{i=0}^{T+T'} \varphi(a_{3i} - c_{3i} + \mu_{3i}), \end{aligned}$$

where $\boldsymbol{\mu}_3$ and \mathbf{a}_3 are given in (4.3.2) and \mathbf{c}_3 is given in (4.3.3). The second probability in the right-hand side of (4.6.5) is $\prod_{i=1}^{T+T'} \varphi(x_i - x_{i+1}) dx_{i+1}$. We finish the proof by collating all terms and noting

$$\prod_{i=0}^{T+T'} \varphi(a_{3i} - c_{3i} + \mu_{3i}) = \prod_{i=0}^{T+T'} \varphi(x_i - x_{i+1}).$$

□

4.6.3 Proof of Theorem 4.4.1

The proof of Theorem 4.4.1 is similar to the proof of Theorem 4.3.1. We modify the event Ω as follows:

$$\begin{aligned} \Omega = \{ & W(t) < W(t+1) + a + bt < W(t+2) + 2a + b + bt + b't < \\ & W(t+3) + 3a + 2b + b' + (b + b' + b'')t \text{ for all } t \in [0, 1] \}. \end{aligned}$$

By the law of total probability,

$$\begin{aligned} F_{a,b,b',b''}(3|x) &= \int \cdots \int \Pr\{\Omega | W(0)=x_0, \dots, W(4)=x_4, W(0)-W(1)=x\} \\ &\quad \times \Pr\{W(0) \in dx_0, \dots, W(4) \in dx_4 | W(0)-W(1)=x\}. \end{aligned} \quad (4.6.6)$$

Define individually the following processes:

$$\begin{aligned} W_0(t) &= W(t) \\ W_1(t) &= a + bt + W(t+1) \\ W_2(t) &= 2a + b + (b + b')t + W(t+2) \\ W_3(t) &= 3a + 2b + b' + (b + b' + b'')t + W(t+3) \end{aligned}$$

with $0 \leq t \leq 1$ for all processes. The event Ω can be re-written as

$$\Omega = \{W_0(t) < W_1(t) < W_2(t) < W_3(t) \text{ for all } t \in [0, 1]\}.$$

The conditioning introduced in (4.6.6) results in:

$$\begin{aligned}
W_0(0) &= 0 & W_0(1) &= x_1 \\
W_1(0) &= a + x_1 & W_1(1) &= a + b + x_2 \\
W_2(0) &= 2a + b + x_2 & W_2(1) &= 2a + 2b + b' + x_3 \\
W_3(0) &= 3a + 2b + b' + x_3 & W_3(1) &= 3a + 3b + 2b' + b'' + x_4.
\end{aligned}$$

From this, we can express (4.6.6) as

$$\begin{aligned}
& F_{a,b,b',b''}(3|x) \\
&= \int \cdots \int \Pr\{\Omega \mid W_0(0) = 0, \dots, W_3(0) = 3a + 2b + b' + x_3, W_0(1) = x_1, \dots, \\
&\quad W_3(1) = 3a + 3b + 2b' + b'' + x_4, W_0(0) - W_0(1) = x\} \\
&\quad \times \Pr\{W(0) \in dx_0, \dots, W(4) \in dx_4 \mid W(0) - W(1) = x\}. \tag{4.6.7}
\end{aligned}$$

The region of integration for (4.6.7) is determined from the following inequalities (see proof of (4.2.5) for similar discussion):

$$x_1 < x_2 + a + b < x_3 + 2a + 2b + b' < x_4 + 3a + 3b + 2b' + b''.$$

Thus, the upper limit of integration is infinity for all x_i . For integration with respect to x_4 , the lower limit is $x_3 - a - b - b' - b''$. For integration with respect x_3 , the lower limit is $x_2 - a - b - b'$. Finally, for x_2 , the lower limit is $x_1 - a - b = -x - a - b$. Now using (4.2.3) with $n = 3$ we obtain

$$\begin{aligned}
& \Pr\{\Omega \mid W_0(0) = 0, \dots, W_3(0) = 3a + 2b + b' + x_3 \\
&\quad W_0(1) = x_1, \dots, W_3(1) = 3a + 3b + 2b' + b'' + x_4, W_0(0) - W_0(1) = x\} \\
&= \exp(-\|\boldsymbol{\mu}_4\|^2/2 + \boldsymbol{\mu}'_4(\mathbf{c}_4 - \mathbf{a}_4)) \det[\varphi(a_{4i}, c_{4j})]_{i,j=0}^3 / \prod_{i=0}^3 \varphi(a_{4i} - c_{4i} + \mu_{4i}),
\end{aligned}$$

$\boldsymbol{\mu}_4$, \mathbf{a}_4 and \mathbf{c}_4 are given in (4.4.2). The second probability in the right-hand side of (4.6.7) is $\prod_{i=1}^3 \varphi(x_i - x_{i+1}) dx_{i+1}$. Using the fact

$$\prod_{i=0}^3 \varphi(a_{4i} - c_{4i} + \mu_{4i}) = \prod_{i=0}^3 \varphi(x_i - x_{i+1}),$$

and collecting all results we complete the proof. \square

Chapter 5

Power of the MOSUM test for online detection of a transient change in mean

Abstract: In this chapter we discuss an on-line moving sum (MOSUM) test for detection of a transient change in the mean of a sequence of i.i.d. normal random variables. By using newly developed theory for a continuous time Gaussian process derived in Chapter 4, and subsequently correcting the results for discrete time, accurate approximations for the Average Run Length (ARL) and power of the test are provided. We check theoretical results against simulations, compare the power of the MOSUM test with that of the CUSUM and briefly consider the cases of non-normal r.v.'s and weighted sums. The content of this chapter has been published in [84].

5.1 Introduction: Statement of the problem

This chapter considers the discrete-time analogue of the continuous-time change-point detection problem studied in Section 4.5 of Chapter 4. The results of the chapter will utilise many of the boundary crossing probabilities derived in Chapter 4 after subsequently correcting the expressions for discreteness. Let us formulate the change-point problem in discrete time.

Suppose one sequentially observes the i.i.d. normal random variables $\varepsilon_1, \varepsilon_2 \dots$ with known mean μ and variance σ^2 . At some unknown change-point ν , the random variables ε_i ($\nu + 1 \leq i \leq \nu + l$) see a change in mean to $\mu + A$ for some positive A and then the pre-change state is resumed. The goal of any detection procedure for this type of change is to detect the change with high probability, subject to a false alarm constraint. Formatted under the hypothesis testing framework, the null hypothesis is $\mathbb{H}_0: \nu = \infty$ and hence $\mathbb{E}\varepsilon_j = \mu$ for all $j = 1, 2, \dots$. The alternative hypothesis is $\mathbb{H}_1: \nu < \infty$ and therefore

$$\mathbb{H}_1 : \mathbb{E}\varepsilon_j = \begin{cases} \mu & \text{if } j \leq \nu \text{ or } j > \nu + l \\ \mu + A & \text{if } \nu < j \leq \nu + l, \end{cases} \quad (5.1.1)$$

with $j = 1, 2, \dots$. Under \mathbb{H}_1 , the arrival time of the signal is $\nu + 1$ (it is unknown), the length of the signal is l (it can be known or unknown) and the amplitude is $A > 0$ (it can be known or unknown). Figure 5.1 displays the values $\mathbb{E}\varepsilon_j$ under \mathbb{H}_1 .

In this chapter, we study characteristics of an on-line moving sum (MOSUM) test for detecting this transient change. This test is discussed comprehensively in Chapter 1. For a fixed sample size, an excellent survey of several tests related to the off-line MOSUM procedure also designed to detect a transient change in the mean value of a sequence of normal random variables is provided in [141]. Here the author considers tests where the length of the transient change is not known in advance and hence the test statistic depends on the maximum over all potential lengths of transient change and all possible locations, see discussions in Section 1.3.2.3 in Chapter 1 and also [140]. Also for the offline regime only, the problem of testing for the existence of a multiple changes in the means of i.i.d. random variables has been studied in [23]. Here, a MOSUM like statistic is used for detecting any possible number of change-points in a sample of fixed length and the values of the mean after the change-point do not necessarily have to be known. This can be seen as a generalisation of the problem considered in this chapter, if one considers only the offline setting. When considering the power of the test, it was shown in Theorem 2.2 of [23] that the MOSUM variant test rejects with asymptotic power one. We refer to Section 1.3.2.3 of Chapter 1 for a brief introduction to statistic studied in [23].

The MOSUM test applied for testing a change in mean of a sequence of i.i.d. r.v.'s was considered in [5] in the off-line regime. Here, the authors consider the cases when μ is known or unknown and σ^2 is known or unknown. In [6], application of MOSUM with squared residuals was used to detect a change in variance. In both papers, the critical values for the test were approximated unsatisfactorily by ignoring correlations between the moving sums. For a fixed sample size, the asymptotic critical values for the off-line MOSUM test were correctly obtained in [18] but there was no attempt made to correct results for discrete time; this idea was considered in Chapter 2. Moreover, to the best of the authors knowledge, only Monte Carlo methods have been used to approximate the power of the MOSUM test and no explicit approximations have been developed. The multivariate extension of this change-point when the length of transient change is infinitely long is discussed in [16]. Here the authors propose a new method for the online multivariate change-point problem and compare against the limited number of existing procedures.

The chapter is structured as follows. In Section 5.2 we formulate the MOSUM test and define the characteristics of interest: ARL (average run length) and the power of the test. In Section 5.3 we briefly recall previous results on approximations for ARL. Section 5.4 forms the bulk of this chapter. In this section we develop approximations for the power of the test for different cases depending on the relation between l and the width of the chosen MOSUM window. In Section 5.5, we study the power of the MOSUM test and compare it with the power of the celebrated CUSUM test. Here we will investigate the power as a function of l/L where L is the window length associated with the MOSUM test (to be formulated shortly). The purpose is to investigate the loss of power when the exact value of l is not known, and L is misspecified. Finally, in Section 5.6 we assess the accuracy of approximations when

the original r.v.'s are not normal and the weights are not uniform.

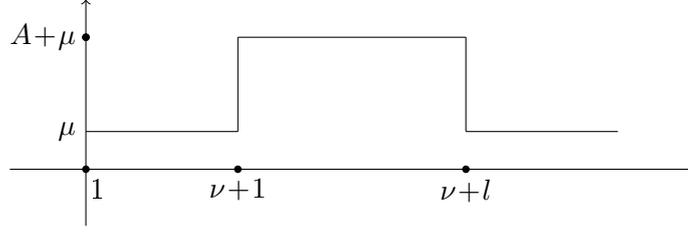


Figure 5.1: Means of r.v. ε_j under the alternative.

5.2 Formulation of MOSUM statistic and its characteristics

5.2.1 Definition of MOSUM test

Let P_i and \mathbb{E}_i denote probability and expectation under \mathbb{H}_i for $i = 0, 1$, respectively. For a fixed positive integer L , define the moving sums

$$S_{n,L} := \sum_{j=n+1}^{n+L} \varepsilon_j \quad (n = 0, 1, \dots). \quad (5.2.1)$$

As demonstrated in Section 1.3.2.1, the MOSUM test is motivated by the log-likelihood ratio test and is defined as follows. After sequentially observing $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{L-1}$ without actioning, the MOSUM stopping rule is

$$\tau(H) := \inf \left\{ n \geq L : \max_{0 \leq k \leq n-L} S_{k,L} \geq H \right\}, \quad (5.2.2)$$

where H is a threshold suitably chosen to satisfy a false alarm constraint; in this case, the ARL constraint $\mathbb{E}_0 \tau(H) = C$ for some (usually large) pre-specified constant C .

Instead of $\tau(H)$, it is sometimes more convenient to use form outlined in (1.3.10) of Chapter 1. By defining

$$\tau_S(H) := \inf \{ n \geq 0 : S_{n,L} \geq H \}, \quad (5.2.3)$$

then the relation between $\tau(H)$ and $\tau_S(H)$ is simply

$$\tau(H) = \tau_S(H) + L. \quad (5.2.4)$$

5.2.2 Main characteristics of interest

In this chapter we shall discuss and derive approximations for the following two characteristics of the test (5.2.2):

1. The ARL to false alarm: $\mathbb{E}_0 \tau(H) = \mathbb{E}_0 \tau_S(H) + L$.

2. The power of the test: For $\nu' = \nu - L$,

$$\mathcal{P}_S(H, A, L) := \lim_{\nu \rightarrow \infty} \mathbb{P}_1\{S_{n,L} > H \text{ for some } n \in [\nu' + 1, \nu + l - 1] \mid \tau_S(H) > \nu'\}. \quad (5.2.5)$$

Formally, we require $\nu \rightarrow \infty$ in (5.2.5). This is to ensure that the sequence of moving sums $\{S_{n,L}\}_n$ reaches the stationary behaviour under the null hypothesis and given that we have not crossed the threshold H . However, as discussed below in Section 5.4.4.3, this stationary regime is reached very quickly and in all approximations below it is enough to only require $\nu \geq 2L$.

5.2.3 Standardisation of MOSUM statistics

It will be more convenient to use the standardised moving sums $S_{n,L}$. Under \mathbb{H}_0 , $\mathbb{E}_0 S_{n,L} = \mu L$ and $\text{Var}_0(S_{n,L}) = \sigma^2 L$, and we define the standardised moving sum process

$$\xi_{n,L} := \frac{S_{n,L} - \mathbb{E}_0 S_{n,L}}{\sqrt{\text{Var}_0(S_{n,L})}} = \frac{S_{n,L} - \mu L}{\sigma \sqrt{L}}, \quad n = 0, 1, \dots$$

Define

$$h = \frac{H - \mu L}{\sigma \sqrt{L}} \text{ so that } H = \mu L + \sigma h \sqrt{L}. \quad (5.2.6)$$

The stopping time

$$\tau_\xi(h) := \inf\{n \geq 0 : \xi_{n,L} \geq h\} \quad (5.2.7)$$

is equivalent to $\tau_S(H)$; that is $\tau_\xi(h) = \tau_S(H)$ and hence $\mathbb{E}_0 \tau_\xi(h) = \mathbb{E}_0 \tau_S(H)$. Moreover, by defining

$$\mathcal{P}_\xi(h, A, L) := \lim_{\nu \rightarrow \infty} \mathbb{P}_1\{\xi_{n,L} > h \text{ for some } n \in [\nu' + 1, \nu + l - 1] \mid \tau_\xi(h) > \nu'\}, \quad (5.2.8)$$

we have $\mathcal{P}_\xi(h, A, L) = \mathcal{P}_S(H, A, L)$.

5.3 Approximation for ARL, the average run length to false alarm

As discussed comprehensively in Chapter 1, see Section 1.2.1 as an example, the threshold H in (5.2.2) is chosen so that $\mathbb{E}_0 \tau(H) = C$ for some (usually large) pre-specified constant $C > 0$. In view of the relations (5.2.4) and (5.2.7), this corresponds to choosing h such that

$$\mathbb{E}_0 \tau_\xi(h) = C - L. \quad (5.3.1)$$

Approximating the quantity $\mathbb{E}_0 \tau_\xi(h)$ was a focus of study in Section 2.8 of Chapter 2. Here we recall one of the main results and refer to Chapter 2 for details of its

derivation.

Approximation 0. From (2.8.2), by setting $h_L = h + 0.8239/\sqrt{L}$:

$$\mathbb{E}_0\tau_\xi(h) \cong -\frac{L \cdot F_2(h; L)}{\theta_L(h)^2 \log(\theta_L(h))}, \text{ with } \theta_L(h) = \frac{F_2(h; L)}{F_1(h; L)},$$

$$F_1(h; L) = \Phi(h)\Phi(h_L) - \varphi(h_L)[h\Phi(h) + \varphi(h)],$$

and

$$\begin{aligned} F_2(h; L) &= \frac{\varphi^2(h_L)}{2} [(h^2 - 1 + \sqrt{\pi}h)\Phi(h) + (h + \sqrt{\pi})\varphi(h)] \\ &\quad - \varphi(h_L)\Phi(h_L) [(h + h_L)\Phi(h) + \varphi(h)] + \Phi(h)\Phi^2(h_L) \\ &\quad + \int_0^\infty \Phi(h-y) [\varphi(h_L + y)\Phi(h_L - y) - \sqrt{\pi}\varphi^2(h_L)\Phi(\sqrt{2}y)] dy. \end{aligned}$$

Here, $\varphi(x)$ and $\Phi(x)$ are the standard normal density and distribution functions respectively and are obtained by setting $\theta = 1$ in the following

$$\varphi_\theta(x) := \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/(2\theta)}, \quad \Phi_\theta(t) := \int_{-\infty}^t \varphi_\theta(x) dx. \quad (5.3.2)$$

As comprehensively shown in Section 2.8 of Chapter 2, Approximation 0 is very accurate. For example, with $h = 3$ and $L = 10$, Approximation 0 yields $\mathbb{E}_0\tau_S(h) \cong 1551$ whereas Monte Carlo simulations with a sample size of 100,000 yield 1550 ± 1 . The author refers the reader to Chapter 2 for a deeper assessment of the accuracy of Approximation 0. In the next section we will develop approximations for the power of the MOSUM test. These will be developed in a similar manner to how $\mathbb{E}_0\tau_\xi(h)$ was approximated in the sense that we approximate a discrete-time problem with a continuous-time problem and subsequently correct results for discrete time (see Chapter 2).

5.4 Power of the test

5.4.1 Equivalent definition for the power of the test $\mathcal{P}_\xi(h, A, L)$

Assume \mathbb{H}_1 so that $\nu < \infty$, and that ν is suitably large. Recall $\nu' = \nu - L$. If the barrier H is reached by the test (5.2.3) for any sum $S_{n,L}$ with $n \leq \nu'$ then, since there are no parts of the signal in the sums $S_{0,L}, \dots, S_{\nu',L}$, we classify the event of reaching the barrier as a false alarm. Each one of the sums $S_{\nu'+1,L}, \dots, S_{\nu+l-1,L}$ has mean larger than $L\mu$ as it contains at least a part of the signal. Reaching the barrier H by any of these sums will be classified as a correct detection of the signal. If neither of these sums reaches H , then we say that we failed to detect the signal and further events when $S_{n,L} \geq H$ with $n \geq \nu + l$ will again be classified as false alarms. In Figure 5.2 we display the values $\mathbb{E}_1 S_{n,L}$ as a function of n .

In this chapter, from the definition of $\mathcal{P}_S(H, A, L)$, which is given in (5.2.5), we are interested in detecting a change when ν occurs in the distant future ($\nu \rightarrow \infty$)

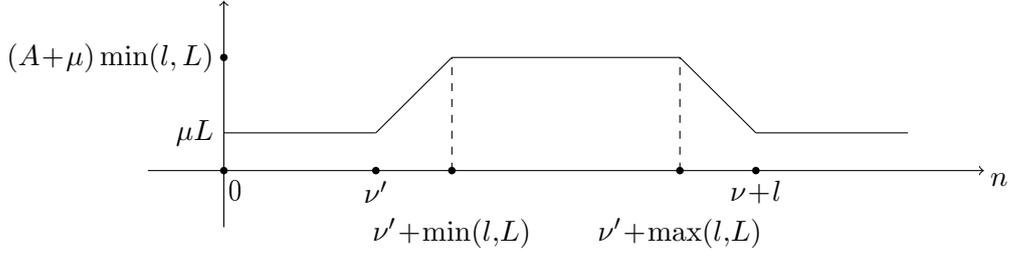


Figure 5.2: $\mathbb{E}_1 S_{n,L}$ as a function of n

given a false alarm has not been raised up to its arrival; this is to ensure the process $\{S_{n,L}\}_n$ reaches the stationary state before the change occurs. In some sense, letting $\nu \rightarrow \infty$ is excessive. The results developed in this chapter are more general and can be applied for all $\nu > 2L$; this is due to the fact that the stationary behaviour of $\{S_{n,L}\}_n$ is achieved very quickly.

Define the function

$$Q(n; A, L, \nu') := \begin{cases} 0 & \text{for } n \leq \nu' \text{ or } n \geq \nu + l \\ A(n - \nu') & \text{for } \nu' < n \leq \nu' + \min(l, L) \\ A \min(l, L) & \text{for } \nu' + \min(l, L) < n \leq \nu' + \max(l, L) \\ A(L + l + \nu' - n) & \text{for } \nu' + \max(l, L) < n \leq \nu + l - 1. \end{cases}$$

By subtracting $\mathbb{E}_1 S_{n,L}$ from the threshold H , the power of the test given in (5.2.5) can be expressed in terms of probability under \mathbb{H}_0 :

$$\mathcal{P}_S(H, A, L) = \lim_{\nu \rightarrow \infty} \mathbb{P}_0 \{ S_{n,L} > H - Q(n; A, L, \nu') \text{ for some } n \in [\nu' + 1, \nu + l - 1] \mid \tau_S(H) > \nu' \}.$$

After standardisation of $S_{n,L}$, the equivalent definition of the power is

$$\mathcal{P}_\xi(h, A, L) = \lim_{\nu \rightarrow \infty} \mathbb{P}_0 \left\{ \xi_{n,L} > h - \frac{Q(n; A, L, \nu')}{\sigma \sqrt{L}} \text{ for some } n \in [\nu' + 1, \nu + l - 1] \mid \tau_\xi(h) > \nu' \right\}, \quad (5.4.1)$$

where we once again recall $\mathcal{P}_S(H, A, L) = \mathcal{P}_\xi(h, A, L)$. To approximate $\mathcal{P}_\xi(h, A, L)$, we shall firstly approximate the problem in the continuous-time setting by deriving what we shall call the diffusion approximation. We will then correct the diffusion approximation for discrete time.

5.4.2 Continuous-time (diffusion) approximation for $\mathcal{P}_\xi(h, A, L)$

For deriving the diffusion approximation of (5.4.1) we replace the discrete time process $\xi_{0,L}, \dots, \xi_{\nu+l,L}$ with a continuous time process $S(t)$, $t \in [0, T]$ with $T = (\nu + l)/L$, as follows.

Set $\Delta = 1/L$ and define $t_n = n\Delta \in [0, T]$ with $n = 0, 1, \dots, \nu + l$. Define a piece-wise linear continuous-time process $S_t^{(L)}$, $t \in [0, T]$:

$$S_t^{(L)} = \frac{1}{\Delta} [(t_n - t)\xi_{n-1,L} + (t - t_{n-1})\xi_{n,L}] \quad \text{for } t \in [t_{n-1}, t_n], \quad n = 1, \dots, \nu + l.$$

By construction, the process $S_t^{(L)}$ is such that $S_{t_n}^{(L)} = \xi_{n,L}$ for $n = 0, \dots, \nu + l$. Fix $\gamma = A\sqrt{L}/\sigma$, $\kappa = \nu'/L$, $\lambda = l/L$ and define the function

$$Q(t; \gamma, \kappa, \lambda) = \begin{cases} 0 & \text{for } t \leq \kappa \text{ or } t \geq \kappa + 1 + \lambda. \\ \gamma(t - \kappa) & \text{for } \kappa < t \leq \kappa + \min(1, \lambda) \\ \gamma \min(1, \lambda) & \text{for } \kappa + \min(1, \lambda) < t \leq \kappa + \max(1, \lambda) \\ \gamma(1 + \lambda + \kappa - t) & \text{for } \kappa + \max(1, \lambda) < t \leq \kappa + 1 + \lambda. \end{cases}$$

Lemma 5.4.1 *Let $A = A(L) = \gamma\sigma/\sqrt{L}$ for some $\gamma > 0$ and assume $L \rightarrow \infty$. Then under \mathbb{H}_1 the limiting process $S(t) = \lim_{L \rightarrow \infty} S_t^{(L)}$, where $t \in [0, T]$, is a Gaussian process with marginal distributions $S(t) \sim N(Q(t; \gamma, \kappa, \lambda), 1)$ for all $t \in [0, T]$ and autocorrelation function $R_S(t, t + s) = R(s) = \max\{0, 1 - |s|\}$. Under \mathbb{H}_0 , the limiting process $S(t)$ is a Gaussian second-order stationary process with marginal distributions $S(t) \sim N(0, 1)$ for all $t \in [0, T]$ and the same autocorrelation function $R(t)$.*

The Gaussian process $S(t)$ with zero mean and autocorrelation function $R(s) = \max\{0, 1 - |s|\}$ is referred to as the Slepian process. The Slepian process $S(t)$ can be expressed in terms of the standard Brownian motion process $W(t)$ as follows:

$$S(t) = W(t) - W(t + 1), \text{ for } t \geq 0. \quad (5.4.2)$$

Values of $\mathbb{E}_1 S(t)$ are shown in Figure 5.3.

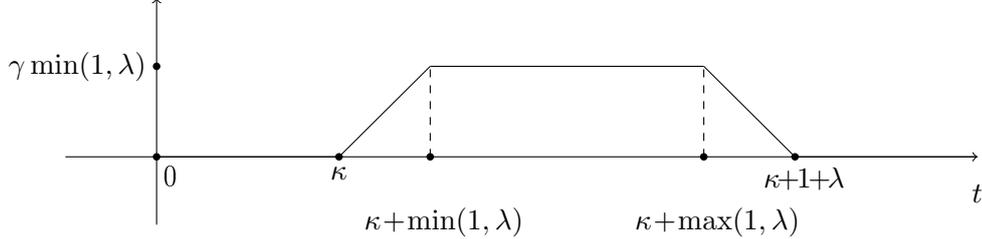


Figure 5.3: $\mathbb{E}_1 S(t)$ as a function of t

The diffusion approximation for the power of the test is

$$\mathcal{P}(h, A) := \quad (5.4.3)$$

$$\lim_{\kappa \rightarrow \infty} P_0\{S(t) > h - Q(t; \gamma, \kappa, \lambda) \text{ for some } n \in [\kappa, \kappa + 1 + \lambda] \mid \tilde{\tau}(h) > \kappa\},$$

where $\tilde{\tau}(h) = \inf\{t > 0 : S(t) > h\}$. That is, we make the approximation

$$\mathcal{P}_\xi(h, A, L) \cong \mathcal{P}(h, A) \quad (5.4.4)$$

by assuming $L \rightarrow \infty$. The diffusion approximation should provide fairly accurate approximations if L is very large. It should, however, be corrected for discrete time for smaller L . The complexity of computation of the diffusion approximation $\mathcal{P}(h, A)$ and its discrete-time corrected version depends on the choice of L in comparison to l . We consider three different cases which are distinguished by the value of $\lambda = l/L$: (a) $\lambda = 1$, (b) $\lambda > 1$ and (c) $\lambda < 1$.

5.4.3 Case (a): $\lambda = 1$

In this section, we consider the case when $L = l$ or, equivalently, $\lambda = 1$. In this case, the diffusion approximation for $\mathcal{P}_\xi(h, A, L)$ given in (5.4.3) reduces to

$$\mathcal{P}(h, A) = \lim_{\kappa \rightarrow \infty} \mathbb{P}_0 \{S(t) \geq h - Q(t; \gamma, \kappa) \text{ for some } t \in [\kappa, \kappa + 2] \mid \tilde{\tau}(h) > \kappa\}, \quad (5.4.5)$$

where $Q(t; \gamma, \kappa) = \gamma \max\{0, 1 - |t - (\kappa + 1)|\}$. The barrier $h - Q(t; \gamma, \kappa)$ is depicted in Figure 5.4.

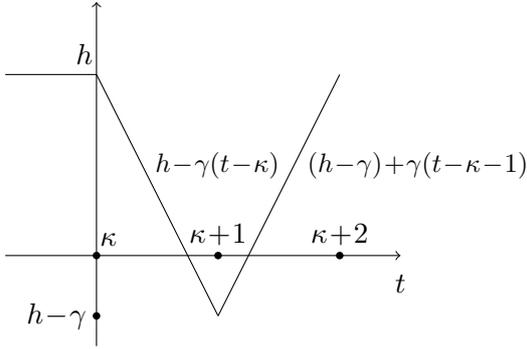


Figure 5.4: Barrier $h - Q(t; \gamma, \kappa)$ for $\lambda = 1$.

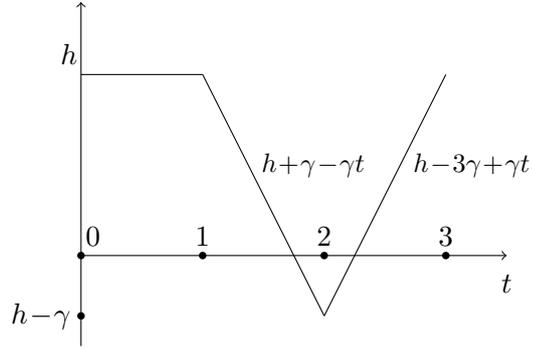


Figure 5.5: Barrier $B(t; h, 0, -\gamma, \gamma)$.

5.4.3.1 Approximation for the diffusion approximation

The probability (5.4.5) was considered in of Section 4.5 Chapter 4, where approximations accurate to more than 4 decimal places were developed. Define the following two conditional probabilities:

$$F_{h,0}(1|x) := \mathbb{P}_0(S(t) < h \text{ for all } t \in [0, 1] \mid S(0) = x),$$

$$F_{h,0,-\gamma,\gamma}(3|x) := \mathbb{P}_0(S(t) < B(t; h, 0, -\gamma, \gamma) \text{ for all } t \in [0, 3] \mid S(0) = x),$$

where the barrier $B(t; h, 0, -\gamma, \gamma)$ is defined as

$$B(t; h, 0, -\gamma, \gamma) = \begin{cases} h, & 0 \leq t \leq 1 \\ h - \gamma(t - 1), & 1 < t \leq 2 \\ h - \gamma + \gamma(t - 2), & 2 < t \leq 3 \\ 0 & \text{otherwise,} \end{cases}$$

and is depicted in Figure 5.5. From (4.5.7) in Chapter 4 (and [145]) we obtain

$$\mathcal{P}(h, A) \cong 1 - \frac{F_{h,0,-\gamma,\gamma}(3|0)}{F_{h,0}(1|0)}, \quad (5.4.6)$$

where

$$F_{h,0}(1|x) = \Phi(h) - \exp(-(h^2 - x^2)/2) \Phi(x) \quad (5.4.7)$$

and

$$F_{h,0,-\gamma,\gamma}(3|x) = \frac{e^{\gamma^2/2}}{\varphi(x)} \int_{-x-h}^{\infty} \int_{x_2-h+\gamma}^{\infty} e^{-\gamma(x_3-x_2)} dx_3 dx_2 \times$$

$$\det \begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h+\gamma) & \Phi(-x_3-2h+\gamma) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h+\gamma) & \Phi(-x-x_3-h+\gamma) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3+\gamma) & \Phi(x_2-x_3+\gamma) \\ \varphi(x_3+3h-\gamma+x) & \varphi(x_3+2h-\gamma-x_2) & \varphi(h) & \Phi(h) \end{bmatrix}.$$

To compute the approximation (5.4.6) one needs to numerically evaluate a two-dimensional integral which is a routine problem for modern computers. Recall that in this thesis we use the convention that $\Phi(x)$ is explicit and not an integral. This is because $\Phi(x)$ can be easily evaluated by all statistical software.

5.4.3.2 Correcting the diffusion approximation (5.4.6) for discrete time

To correct (5.4.6) for discrete time, we must correct the continuous-time probabilities $F_{h,0}(1|x)$ and $F_{h,0,-\gamma,\gamma}(3|x)$ for discrete time. In Section 2.4.5 of Chapter 2 it is explained that correcting the probability $F_{h,0}(1|x)$ for discrete time amounts to replacing the threshold h by $h_L := h + \omega_L$, where $\omega_L = 0.8239/\sqrt{L}$. Here we shall show that correcting $F_{h,0,-\gamma,\gamma}(3|x)$ for discrete time can be performed in the same manner.

Let $W(t)$ be the standard Brownian Motion process on $[0, \infty)$ with $W(0) = 0$ and $\mathbb{E}W(t)W(s) = \min(t, s)$. Define the event

$$\Omega = \{W(t) - W(t+1) < h, W(t+1) - W(t+2) < h - \gamma t, \\ W(t+2) - W(t+3) < h - \gamma + \gamma t, \forall 0 \leq t \leq 1\}.$$

If $W(i) = x_i$, $i = 0, 1, \dots, 4$, we obtain from the proof of Theorem 4.4.1 in Chapter 4:

$$F_{h,0,-\gamma,\gamma}(3|x) = \int \dots \int \Pr\{\Omega \mid W(i) = x_i, i = 0, 1, 2, 3, 4, W(0) = 0, W(0) - W(1) = x\} \\ \times \Pr\{W(i) \in dx_i, i = 0, 1, 2, 3, 4 \mid W(0) = 0, W(0) - W(1) = x\}. \quad (5.4.8)$$

It follows from the proof of (5.4.8) that correcting for discrete time amounts to correcting the following probability for discrete time

$$\Pr\{\Omega \mid W(i) = x_i, i = 0, 1, 2, \dots, 4, W(0) = 0, W(0) - W(1) = x\} \\ = \Pr\{\sqrt{2}W_1(t) < h, \sqrt{2}W_2(t) < h - \gamma t, \sqrt{2}W_3(t) < h - \gamma + \gamma t, \forall 0 \leq t \leq 1 \mid \\ W(i) = x_i, i = 0, 1, 2, \dots, 4, W(0) = 0, W(0) - W(1) = x\}, \quad (5.4.9)$$

where $W_i(t) = \frac{\sqrt{2}}{2}[W(t+i-1) - W(t+i)]$, $i = 1, 2, 3$. Due to the conditioning on the rhs of (5.4.9), the processes $W_i(t)$ can be treated as independent Brownian motion processes. In view of the fact that Brownian motion has independent increments,

correcting formula (5.4.8) for discrete time is equivalent to correcting the probabilities $\Pr(\sqrt{2}W(t) < h + bt, \forall 0 \leq t \leq 1)$ for discrete time, where $b \in \{0, -\gamma, \gamma\}$. Correction of $\Pr(\sqrt{2}W(t) < h + bt, \forall 0 \leq t \leq 1)$ for discrete time is discussed in Section 2.4.5.2 of Chapter 2. In particular, it is shown in Section 2.4.5.2 that the expected excess over the boundary for a discretised form of $\sqrt{2}W(t)$ does not depend on the value of b . Hence, results from Section 2.4.5.2 imply that to correct $F_{h,0,-\gamma,\gamma}(3|x)$ for discrete time, one should simply replace the threshold h by $h_L := h + \omega_L$.

Approximation 1. For $\lambda = 1$, the corrected diffusion approximation for the power of the test is

$$\mathcal{P}_\xi(h, A, L) \cong 1 - \frac{F_{h_L,0,-\gamma,\gamma}(3|0)}{F_{h_L,0}(1|0)}, \text{ where } h_L := h + \omega_L. \quad (5.4.10)$$

5.4.3.3 Simulation studies

In this section, we evaluate the performance of Approximation 1. In Figures 5.6, the thicker black dashed line corresponds to the empirical values of the BCP $\mathcal{P}_\xi(h, A, L)$ computed from 100,000 simulations with different values of L and γ , where $\mu = 0$ and $\sigma = 1$. The solid red line corresponds to Approximation 1. The dot-dashed blue line corresponds to the diffusion approximation given in (5.4.6). The axis are: the x -axis shows the value of γ . The y -axis denotes the probabilities of reaching the barrier. The graphs, therefore, show the empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ and values of approximation (5.4.10).

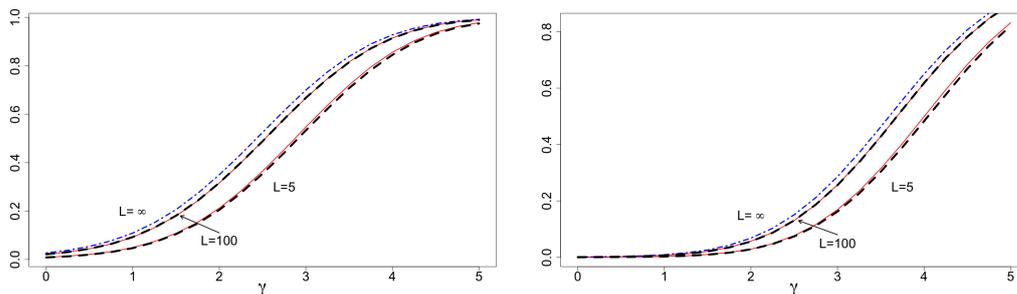


Figure 5.6: Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (thick dashed black) and its Approximation 1 (solid red) for two different values of h . Left: $h = 3$. Right: $h = 4$.

From Figure 5.6, we see that Approximation 1 is very accurate even for a very small $L = 5$. We also see the significance of the discrete-time correction; whilst the diffusion approximation provides sensible results should you compare it with $L = 100$, for $L = 5$ the diffusion approximation is very far off.

5.4.4 Case (b): $\lambda > 1$

For $\lambda > 1$, the diffusion approximation for $\mathcal{P}_\xi(h, A, L)$ given in (5.4.3) reduces to

$$\mathcal{P}(h, A) = \lim_{\kappa \rightarrow \infty} \mathbb{P}_0 \{S(t) \geq h - Q(t; \gamma, \kappa) \text{ for some } t \in [\kappa, \kappa + 1 + \lambda] \mid \tilde{\tau}(h) > \kappa\}, \quad (5.4.11)$$

where

$$Q(t; \gamma, \kappa) = \begin{cases} \gamma(t - \kappa) & \kappa \leq t \leq \kappa + 1 \\ \gamma & \kappa + 1 < t \leq \kappa + \lambda \\ \gamma(\kappa + 1 + \lambda - t) & \kappa + \lambda < t \leq \kappa + 1 + \lambda \\ 0 & \text{otherwise.} \end{cases}$$

The barrier $h - Q(t; \gamma, \kappa)$ is depicted in Figure 5.7. In this section we only consider $1 < \lambda \leq 2$ since larger values of λ result in ungainly high-dimensional integrals. The methodology, however, could still be applied if one has the patience to evaluate integrals of dimension five or higher accurately.

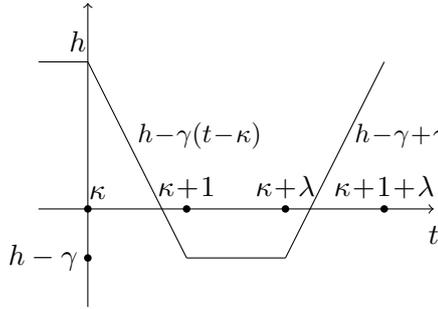


Figure 5.7: Barrier $h - Q(t; \gamma, \kappa)$ for $\lambda > 1$.

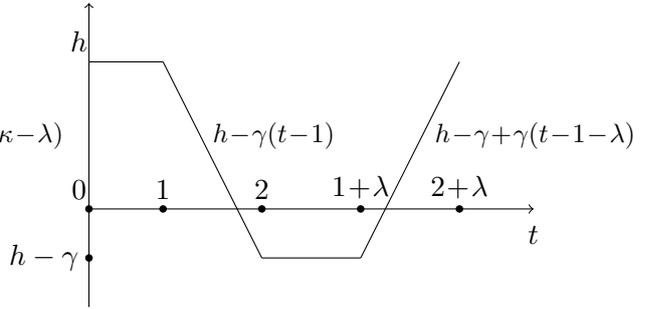


Figure 5.8: Barrier $B(t; h, 0, -\gamma, 0, \gamma)$.

5.4.4.1 Approximating (5.4.11) when $\lambda = 2$

To approximate (5.4.11), one could adopt the same approach in Chapter 4 which led to the form of approximation (5.4.6). Namely, define the conditional probability

$$F_{h,0,-\gamma,0,\gamma}(2 + \lambda | x) := \mathbb{P}_0(S(t) < B(t; h, 0, -\gamma, 0, \gamma) \text{ for all } t \in [0, 2 + \lambda] \mid S(0) = x),$$

where

$$B(t; h, 0, -\gamma, 0, \gamma) = \begin{cases} h, & 0 \leq t \leq 1 \\ h - \gamma(t - 1), & 1 < t \leq 2 \\ h - \gamma, & 2 < t \leq \lambda + 1 \\ h - \gamma + \gamma(t - \lambda - 1), & \lambda + 1 < t \leq 2 + \lambda \\ 0 & \text{otherwise.} \end{cases}$$

The barrier $B(t; h, 0, -\gamma, 0, \gamma)$ is depicted in Figure 5.8. The approximation to (5.4.11) is then

$$\mathcal{P}(h, A) \cong 1 - \frac{F_{h,0,-\gamma,0,\gamma}(2 + \lambda | 0)}{F_{h,0}(1 | 0)}. \quad (5.4.12)$$

Unfortunately, whilst the probability $F_{h,0,-\gamma,0,\gamma}(2 + \lambda|0)$ can be computed, it is often intractable because of the requirement to evaluate a high dimensional integral. For integer λ , a $(\lambda + 1)$ -dimensional integral must be evaluated. For non-integer λ , the situation is harder as a $4\lceil\lambda\rceil$ dimensional integral must be evaluated. The only case of interest where this methodology produces practical formulas is when $\lambda = 2$. The following theorem allows one to evaluate the numerator in (5.4.12). The denominator can be obtained from (5.4.7).

Theorem 5.4.1 *For $x < h$ we have*

$$F_{h,0,-\gamma,0,\gamma}(4|x) = \frac{e^{\gamma^2}}{\varphi(x)} \int_{-x-h}^{\infty} dx_2 \int_{x_2-h+\gamma}^{\infty} dx_3 \int_{x_3-h+\gamma}^{\infty} e^{-\gamma(x_4-x_2)} dx_4 \times \quad (5.4.13)$$

$$\det \begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h+\gamma) & \varphi(-x_4-3h+2\gamma) & \Phi(-x_4-3h+2\gamma) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h+\gamma) & \varphi(-x-x_4-2h+2\gamma) & \Phi(-x-x_4-2h+2\gamma) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3+\gamma) & \varphi(x_2-x_4-h+2\gamma) & \Phi(x_2-x_4-h+2\gamma) \\ \varphi(x_3+3h-\gamma+x) & \varphi(x_3+2h-\gamma-x_2) & \varphi(h) & \varphi(x_3-x_4+\gamma) & \Phi(x_3-x_4+\gamma) \\ \varphi(x_4+4h-2\gamma+x) & \varphi(x_4+3h-2\gamma-x_2) & \varphi(x_4+2h-\gamma-x_3) & \varphi(h) & \Phi(h) \end{bmatrix}.$$

The proof of Theorem 5.4.1 is given in Appendix A in Section 5.7.

5.4.4.2 Correcting the diffusion approximation (5.4.12) for discrete time

To correct approximation (5.4.12) for discrete time, we adopt the methodology of Section 5.4.3.2; exactly the same arguments can be used to justify replacement of the threshold h with h_L . This yields the following.

Approximation 2. *For $\lambda = 2$, the corrected diffusion approximation for the power of the test is*

$$\mathcal{P}_\xi(h, A, L) \cong 1 - \frac{F_{h_L,0,-\gamma,0,\gamma}(4|0)}{F_{h_L,0}(1|0)},$$

where $h_L = h + \omega_L$, $F_{h_L,0,-\gamma,0,\gamma}(4|0)$ can be computed using (5.4.13) and $F_{h_L,0}(1|0)$ can be computed from (5.4.7).

Approximation 2 is very accurate; for simulation results see, in particular, Section 5.4.4.5.

5.4.4.3 Approximating (5.4.11) when $1 < \lambda < 2$

For non-integer $\lambda > 1$ the previous methodology requires the evaluation of a $4\lceil\lambda\rceil$ -dimensional integral which is too demanding. To simplify the evaluation of our approximation to (5.4.11), we treat the process $S(t)$ as if it was Markovian. We keep the restriction $1 < \lambda < 2$ as for $\lambda > 2$, even assuming a Markovian nature cannot reduce the dimension of integration to a reasonable number. Also, as follows from the numerical results discussed in Section 5, the case $1 < \lambda < 2$ is practically much more interesting than the case $\lambda > 2$ as the power of the MOSUM test becomes poor as λ becomes large.

For $a, b \in \mathbb{R}$ and $x < a$, define $F_{a,b}(T|x) := \mathbb{P}_0(S(t) < a + bt \text{ for all } t \in [0, T] \mid S(0) = x)$. Deriving our approximation to $\mathcal{P}(h, A)$ in case $1 < \lambda < 2$ will require the following auxiliary results.

Lemma 5.4.2 For $T = 1$, from (4.2.6) of Chapter 4,

$$\begin{aligned} F_{a,b}(1|x) &= \frac{1}{\varphi(x)} \int_{-\infty}^{a+b} \exp(-b^2/2 + b(b-s_1)) \det \begin{bmatrix} \varphi(x) & \varphi(x+s_1-a-b) \\ \varphi(a) & \varphi(s_1-b) \end{bmatrix} ds_1 \\ &= \Phi(a+b) - \exp(-(a^2-x^2)/2 - b(a-x)) \Phi(x+b). \end{aligned} \quad (5.4.14)$$

From here, it follows

$$P_0(S(1) \in ds_1 | S(t) < a + bt \quad \forall t \in [0, 1], S(0) = x) = f_{a,b}(s_1|x) ds_1,$$

where

$$\begin{aligned} f_{a,b}(s_1|x) &= \frac{e^{b^2/2 - bs_1}}{\varphi(x)} \det \begin{bmatrix} \varphi(x) & \varphi(x+s_1-a-b) \\ \varphi(a) & \varphi(s_1-b) \end{bmatrix} \\ &= \frac{e^{b^2/2 - bs_1}}{\varphi(x)} [\varphi(x)\varphi(s_1-b) - \varphi(a)\varphi(x+s_1-a-b)]. \end{aligned}$$

For $T = \theta$ with $0 < \theta < 1$, we obtain from Theorem 4.2.2 of Chapter 4

$$\begin{aligned} F_{a,b}(\theta|x) &= \frac{e^{\theta b^2/2}}{\varphi(x)} \int_{-\infty}^{\infty} \int_{v_0-a-b\theta}^{\infty} e^{b(v_1+x)} \varphi_{1-\theta}(v_0+x) \\ &\quad \times \det \begin{bmatrix} \varphi_{\theta}(-v_0) & \varphi_{\theta}(-v_1-a-b\theta) \\ \varphi_{\theta}(-x+a-v_0) & \varphi_{\theta}(-x-v_1-b\theta) \end{bmatrix} dv_1 dv_0, \end{aligned}$$

where $\varphi_{\theta}(\cdot)$ is defined in (5.3.2). By making the substitution $s_{\theta} = v_0 - v_1$, we can express the integral above in terms of the random variable $S(\theta)$ to obtain the following lemma.

Lemma 5.4.3 For $T = \theta$ with $0 < \theta < 1$,

$$\begin{aligned} F_{a,b}(\theta|x) &= \frac{e^{\theta b^2/2}}{\varphi(x)} \int_{-\infty}^{\infty} \int_{-\infty}^{a+b\theta} e^{b(v_0-s_{\theta}+x)} \varphi_{1-\theta}(v_0+x) \\ &\quad \times \det \begin{bmatrix} \varphi_{\theta}(-v_0) & \varphi_{\theta}(s_{\theta}-v_0-a-b\theta) \\ \varphi_{\theta}(-x+a-v_0) & \varphi_{\theta}(s_{\theta}-v_0-x-b\theta) \end{bmatrix} ds_{\theta} dv_0. \end{aligned}$$

From here, it follows

$$P_0(S(\theta) \in ds_{\theta} | S(t) < a + bt \quad \forall t \in [0, \theta], S(0) = x) = f_{a,b}^{(\theta)}(s_{\theta}|x) ds_{\theta},$$

where

$$\begin{aligned} f_{a,b}^{(\theta)}(s_{\theta}|x) &= \frac{e^{\theta b^2/2}}{\varphi(x)} \int_{-\infty}^{\infty} e^{b(v_0-s_{\theta}+x)} \varphi_{1-\theta}(v_0+x) \det \begin{bmatrix} \varphi_{\theta}(-v_0) & \varphi_{\theta}(s_{\theta}-v_0-a-b\theta) \\ \varphi_{\theta}(-x+a-v_0) & \varphi_{\theta}(s_{\theta}-v_0-x-b\theta) \end{bmatrix} dv_0 \\ &= \frac{1}{\sqrt{2\pi}\varphi(x)} \left\{ e^{s_{\theta}x/(\theta-2)} \varphi_{\theta(2-\theta)}(s_{\theta}-x) - e^{b(x-a)+a(s_{\theta}+x-a)/\theta} \varphi_{\theta(2-\theta)}(s_{\theta}+x) \right\}. \end{aligned}$$

In [95] and [97], where general Markovian test statistics including CUSUM and Shiryaev-Roberts tests were considered, the distribution of the test statistic under the null hypothesis and under the condition that a false alarm has not been raised for very long time, was called quasi-stationary distribution; below we shall adopt this term for the statistic of moving sums.

To formulate our approximation we require an approximation to the quasi-stationary distribution of $S(t)$. More specifically, we require an approximation to

$$q_h(x) = \frac{d}{dx}Q_h(x), \text{ where } Q_h(x) = \lim_{t \rightarrow \infty} P_0(S(t) < x | \tilde{\tau}(h) > t);$$

recall $\tilde{\tau}(h) = \inf\{t > 0 : S(t) > h\}$. Whilst the exact form of $q_h(x)$ is not known, it is possible to obtain a sequence of approximations increasing in accuracy (but also in computational difficulty):

$$q_h(x) \cong q_h^{(i)}(x) = \frac{d}{dx}P_0(S(i) < x | \tilde{\tau}(h) > i), i = 0, 1, 2, \dots$$

In view of (5.4.2) the process $S(t)$ forgets the past after one unit of time hence quickly reaches the stationary behaviour under the condition $S(t) < h$ for all $t < i$. Therefore the quasi-stationary density function can be well approximated simply by taking $i = 1$. That is,

$$q_h(s) \cong q_h^{(1)}(s) = \frac{\int_{-\infty}^h f_{h,0}(s|x)\varphi(x)dx}{\int_{-\infty}^h \int_{-\infty}^h f_{h,0}(s|x)\varphi(x)dx ds} = \frac{\Phi(h)\varphi(s) - \varphi(h)\Phi(s)}{\Phi^2(h) - \varphi(h)[h\Phi(h) + \varphi(h)]},$$

where $f_{h,0}(s|x)$ is given in Lemma 5.4.3. Taking $i = 2$ would have resulted in a more accurate approximation to $q_h(s)$ at the price of needing to numerically evaluate a one-dimensional integral. Numerical studies show that $q_h^{(2)}(s)$ is extremely close to $q_h^{(1)}(s)$ which has the benefit of being explicitly written, see Figure 5.9. Taking $i = 0$ results in $q_h^{(0)}(s) = \varphi(x)/\Phi(h)$; numerical studies and Figure 5.9 (left) show that for small h this density is far from $q_h^{(1)}(s)$ and $q_h^{(2)}(s)$. However, if h is large, the density $q_h^{(0)}(s)$ gives a perfectly satisfactory approximation to the quasi-stationary density $q_h(s)$; this is seen in Figure 5.9, right.

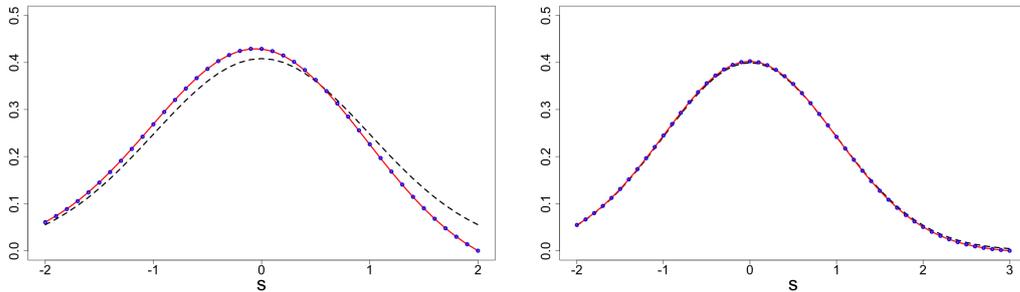


Figure 5.9: The transition densities $q_h^{(0)}(s)$ (dashed black), $q_h^{(1)}(s)$ (solid red) and $q_h^{(2)}(s)$ (dotted blue) for two different values of h . Left: $h = 2$. Right: $h = 3$.

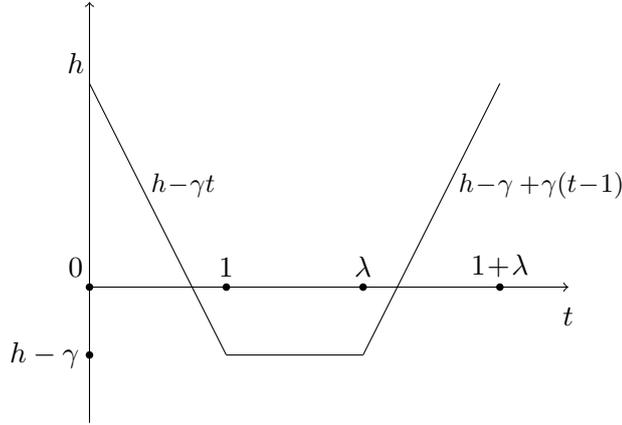


Figure 5.10: Barrier $B_1(t; h, -\gamma, 0, \gamma)$

By treating the process $S(t)$ as if it was Markovian, we can use Lemma 5.4.2 on the interval $[0, 1]$, Lemma 5.4.3 on the interval $[1, \lambda]$ and Lemma 5.4.2 again on the interval $[\lambda, 1 + \lambda]$ to approximate the probability $\mathbb{P}_0(S(t) < B_1(t; h, -\gamma, 0, \gamma))$ for all $t \in [0, 1 + \lambda] \mid S(0) = x$, where the barrier $B_1(t; h, -\gamma, 0, \gamma)$ is defined as

$$B_1(t; h, -\gamma, 0, \gamma) = \begin{cases} h - \gamma t, & 0 < t \leq 1 \\ h - \gamma, & 1 < t \leq 1 + \lambda \\ h - \gamma + \gamma(t - \lambda), & \lambda + 1 < t \leq \lambda + 1 \\ 0 & \text{otherwise,} \end{cases}$$

and is depicted in Figure 5.10. Combining this with the approximation $q_h^{(1)}(s)$ for the quasi-stationary density $q_h(s)$ leads to the following approximation for $\mathcal{P}(h, A)$:

$$\mathcal{P}(h, A) \cong 1 - \int_{-\infty}^h \int_{-\infty}^{h-\gamma} \int_{-\infty}^{h-\gamma} q_h^{(1)}(s_0) f_{h,-\gamma}(s_1 | s_0) f_{h-\gamma,0}^{(\lambda-1)}(s_\lambda | s_1) F_{h-\gamma,\gamma}(1 | s_\lambda) ds_\lambda ds_1 ds_0. \quad (5.4.15)$$

5.4.4.4 Correcting the diffusion approximation (5.4.15) for discrete time

Correcting approximation (5.4.15) for discrete time relies on correcting Lemma 5.4.2, Lemma 5.4.3 and the quasi-stationary density approximation $q_h^{(1)}(s_0)$ for discrete time. We apply exactly the same arguments as in Section 5.4.3.2. As a result, we obtain in the discrete-time corrected forms of $f_{h,-\gamma}^{(1)}(s_1 | s_0)$, $f_{h-\gamma,0}^{(\lambda-1)}(s_\lambda | s_1)$ and $q_h^{(1)}(s_0)$.

Approximation 3. For $1 < \lambda < 2$, the corrected diffusion approximation for the power of the test is obtained by replacing h with $h_L = h + \omega_L$ in (5.4.15).

5.4.4.5 Simulation studies

In this section, we evaluate the performance of Approximations 2 and 3. The style of Figure 5.11 is the same as Figure 5.6 where the solid red line corresponds to the approximations, the dashed black line corresponds to Monte Carlo approximations

for $\mathcal{P}_\xi(h, A, L)$ and the dot-dashed blue line corresponds to the diffusion approximation given in (5.4.15). As in Section 5.4.3.3, empirical values of the BCP $\mathcal{P}_\xi(h, A, L)$ are computed from 100,000 simulations with $\mu = 0$ and $\sigma = 1$.

From Figure 5.11 we see that both Approximations 2 and 3 perform very well.

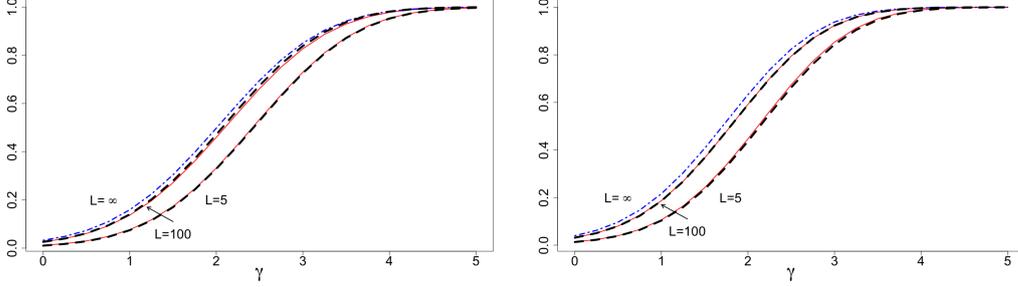


Figure 5.11: Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and its approximations (solid red). Left: $h = 3, \lambda = 1.5$ with Approximation 3. Right: $h = 3, \lambda = 2$ with Approximation 2.

5.4.5 Case (c): $\lambda < 1$

The final case to consider is that of $0 < \lambda < 1$. For $0 < \lambda < 1$, the diffusion approximation for $\mathcal{P}_\xi(h, A, L)$ given in (5.4.3) reduces to

$$\mathcal{P}(h, A) = \lim_{\kappa \rightarrow \infty} \mathbb{P}_0 \{ S(t) \geq h - Q(t; \gamma, \kappa) \text{ for some } t \in [\kappa, \kappa + 1 + \lambda] \mid \tilde{\tau}(h) > \kappa \}, \quad (5.4.16)$$

where

$$Q(t; \gamma, \kappa) = \begin{cases} \gamma(t - \kappa) & \kappa \leq t \leq \kappa + \lambda \\ \gamma\lambda & \kappa + \lambda < t \leq \kappa + 1 \\ \gamma(\kappa + 1 + \lambda - t) & \kappa + 1 < t \leq \kappa + 1 + \lambda \\ 0 & \text{otherwise.} \end{cases}$$

The barrier $h - Q(t; \gamma, \kappa)$ is depicted in Figure 5.12.

5.4.5.1 Approximating (5.4.16) when $0 < \lambda < 1$

For $\lambda \in (0, 1)$, an approximation to $\mathcal{P}(h, A)$ of the form similar to (5.4.6) and (5.4.12) would lead to intractable integrals. As a result, we adopt the same methodology as Section 5.4.4.3, where we form our approximation by treating the process $S(t)$ as if it was Markovian. Repeated application of Lemma 5.4.3 on the intervals $[0, \lambda]$, $[\lambda, 1]$ and $[1, 1 + \lambda]$ allows one to approximate the probability $\mathbb{P}_0(S(t) < B_2(t; h, -\gamma, 0, \gamma) \mid S(0) = x)$, where $B_2(t; h, -\gamma, 0, \gamma)$ is defined by

$$B_2(t; h, -\gamma, 0, \gamma) = \begin{cases} h - \gamma t, & 0 < t \leq \lambda \\ h - \gamma\lambda, & \lambda < t \leq 1 \\ h - \gamma\lambda + \gamma(t - 1), & \lambda + 1 < t \leq \lambda + 1 \\ 0 & \text{otherwise,} \end{cases}$$

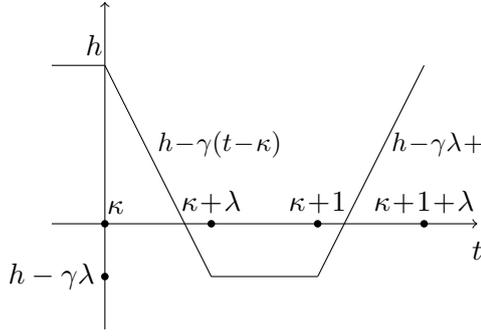


Figure 5.12: Barrier $h - Q(t; \gamma, \kappa)$ for $\lambda < 1$.

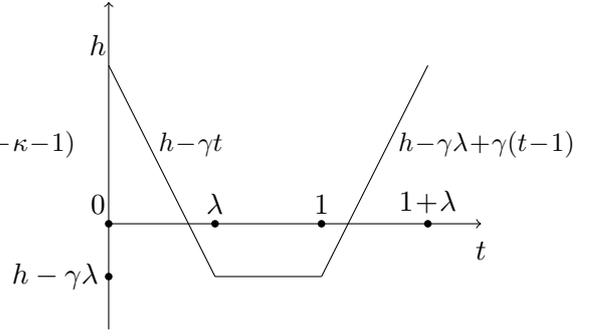


Figure 5.13: Barrier $B_2(t; h, -\gamma, 0, \gamma)$.

and is depicted in Figure 5.13. Combining this with the approximation for the quasi-stationary density leads to the following approximation of $\mathcal{P}(h, A)$:

$$\mathcal{P}(h, A) \cong 1 - \int_{-\infty}^h \int_{-\infty}^{h-\gamma\lambda} \int_{-\infty}^{h-\gamma\lambda} q_h^{(1)}(s_0) f_{h,-\gamma}^{(\lambda)}(s_\lambda | s_0) f_{h-\gamma\lambda,0}^{(1-\lambda)}(s_1 | s_\lambda) F_{h-\gamma\lambda,\gamma}(\lambda | s_1) ds_1 ds_\lambda ds_0. \quad (5.4.17)$$

5.4.5.2 Correcting the diffusion approximation (5.4.17) for discrete time

Correcting (5.4.17) for discrete time can be done in exactly the same manner as for (5.4.15), see Section 5.4.4.4.

Approximation 4. *The corrected diffusion approximation for the power of the test when $0 < \lambda < 1$ can be obtained by replacing h with $h_L = h + \omega_L$ in (5.4.17).*

5.4.5.3 Simulation studies

In this section, we evaluate the performance of Approximation 4. The style of Figure 5.14 is the same as Figures 5.6 and 5.11. For Monte Carlo simulations, we have use 100,000 runs.

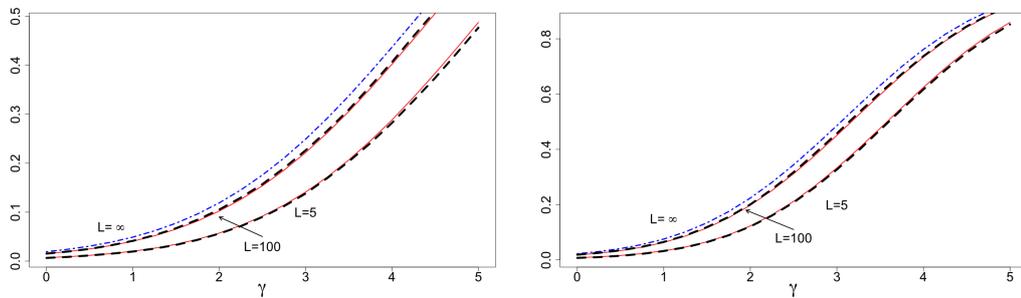


Figure 5.14: Empirical probabilities of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and Approximation 4 (solid red) Left: $h = 3$, $\lambda = 0.5$. Right: $h = 3$, $\lambda = 0.75$.

5.5 Comparison of MOSUM and CUSUM tests

In this section, we draw two comparisons of the MOSUM test. Firstly, we compare the power of the MOSUM test against the well known CUSUM test; a detailed description of the CUSUM test including properties such as ARL can be found in Chapter 1 and [131, Section 8.2]. Secondly, but also simultaneously, we compare the power of the MOSUM test as λ varies in $[0.5, 2]$. The CUSUM test considered in this section assumes the post-change mean A is known. There are variations of the CUSUM test that do not make this assumption, see for example [24], (see also Section 1.3.2.3 in Chapter 1 for variations of the MOSUM test when A is unknown), however these will inevitably have lower power for the case of A known as they utilise less information. The purpose of this section is to make a worst case comparison between the MOSUM process (which does not need to know A to determine its ARL threshold) and the CUSUM test under an ideal scenario with A known. By design, the CUSUM test (for this problem) aims to detect a permanent change in mean rather than a transient change in mean, see Section 1.2 in Chapter 1. We draw comparisons between the two tests to quantify the potential gain in power the MOSUM test has over the CUSUM test by having the additional knowledge that a transient change is present. Furthermore, by studying the power of the MOSUM test as λ varies in $[0.5, 2]$, we investigate when this (potential) advantage is lost.

By defining

$$V_n = \max_{1 \leq k \leq n} \prod_{j=k}^n \exp \left(\frac{(\varepsilon_j - \mu)^2 - (\varepsilon_j - \mu - A)^2}{2\sigma^2} \right),$$

the CUSUM test is

$$\tau_{CS}(H_{CS}) = \inf\{n \geq 1 : V_n > H_{CS}\}, \quad (5.5.1)$$

where H_{CS} is chosen such that $\mathbb{E}_0\tau_{CS}(H_{CS}) = C_{CS}$ for some usually large C_{CS} . The power of the CUSUM test for the transient change considered in this chapter is defined as

$$\mathcal{P}_{CS}(H_{CS}, A) := \lim_{\nu \rightarrow \infty} \mathbb{P}_1\{V_n > H_{CS} \text{ for some } n \in [\nu + 1, \nu + 2l - 1] \mid \tau_{CS}(H_{CS}) > \nu\}.$$

To compare the MOSUM test with the CUSUM test, the threshold H in (5.2.2) and the threshold H_{CS} in (5.5.1) are chosen such that $\mathbb{E}_0\tau(H) = \mathbb{E}_0\tau_{CS}(H_{CS}) = 5000$. Determination of H_{CS} for CUSUM was obtained using tabulated values given in [77, p. 3237] whereas suitable values of H for MOSUM was obtained using Approximation 0 of Section 5.3.

In the first example, we set $A = 1.25$ and $l = 10$. For the MOSUM test, we considered values of $L \in [5, 20]$ to ensure $\lambda \in [0.5, 2]$. For each λ , the values of the approximations developed in this chapter for the power $\mathcal{P}_\xi(h, A, L)$ are displayed in Figure 5.15 (left) with a solid black line. The dot-dashed blue line corresponds to $\mathcal{P}_{CS}(H_{CS}, A)$ which has been obtained via Monte Carlo simulations with 100,000 runs (analytical approximations for the power of CUSUM test are unavailable). In Figure 5.15 (right), we plot the approximations to the power of the MOSUM test

against $\eta = \frac{\max\{1,\lambda\}}{\min\{1,\lambda\}}$. The purpose of such a figure is to compare the power of the test when $L = l/\lambda$ is selected $\eta = 1/\lambda > 1$ times larger than l against the case when L is chosen $\eta = \lambda > 1$ times smaller than l ; $\eta \in [1, 2]$.

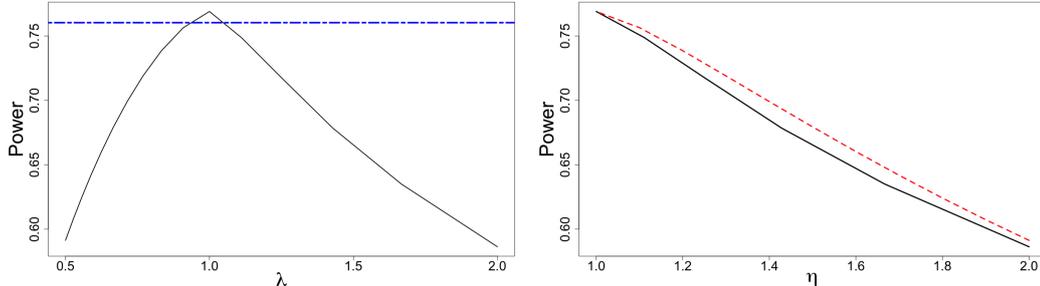


Figure 5.15: Left: Comparison of power between the MOSUM test (solid black) as a function of λ and the CUSUM test (dot-dashed blue) with $A = 1.25$ and $l = 10$. Right: Power of the MOSUM test against η for different λ with $A = 1.25$ and $l = 10$: $\lambda < 1$ (red-dotted) $\lambda > 1$ (solid black).

From Figure 5.15 (left), one can observe the advantage of knowing l since the largest power is obtained for $\lambda = l/L = 1$. For the choice of parameters considered in this example, Figure 5.15 (left) shows that the CUSUM test has power close but still lower than the MOSUM test with $\lambda = 1$. It should be noted, however, that the knowledge of A is required for construction of the CUSUM test but not for the MOSUM test. The values of the approximations in Figure 5.15 (right) suggest the strategy which chooses L which may slightly overestimate (rather than underestimate) l if the exact value of l is not known. For example, for $\lambda = 0.5$, the power of the test is approximately 0.512. The equivalent value with underestimated l is obtained with $\lambda = 2$ and is approximately 0.488. Another sensible strategy for detecting a transient change would be to run several MOSUM tests with different L (and hence H) in parallel; plotting the maximum value of the detection statistics (as a function of L) in case of a change is expected to look like Figures 5.15 (left) and 5.16 (left) distorted by noise.

In the second example, we set $A = 0.5$ and $l = 50$. We considered values of $L \in [25, 100]$ to ensure $\lambda \in [0.5, 2]$. The style of Figure 5.16 is the same as Figure 5.15. For computing $\mathcal{P}_{CS}(H_{CS}, A)$, we have used Monte Carlo simulations with 100,000 runs. The behaviour for the MOSUM test is very similar to the first example considered; the choice of parameters in this example leads to higher power but the general shape, as λ varies, is very similar. More interestingly, in this example the MOSUM test has a significant power advantage over the CUSUM test and this lasts for a wider range of λ when compared to the first example.

5.6 Robustness to non-normality and non-uniform weights

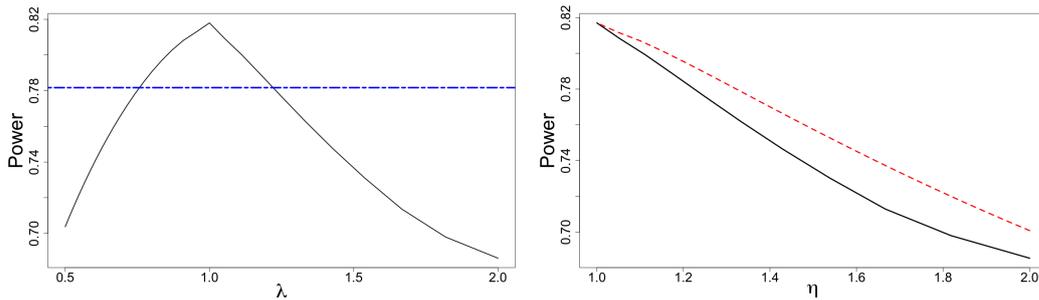


Figure 5.16: Left: Comparison of power between the MOSUM test (solid black) as a function of λ and the CUSUM test (dot-dashed blue) with $A = 0.5$ and $l = 50$. Right: Power of the MOSUM test against η for different λ with $A = 0.5$ and $l = 50$: $\lambda < 1$ (red-dotted) $\lambda > 1$ (solid black).

5.6.1 Robustness to non-normality

Here we assess the quality of approximations developed in this chapter when the original r.v.'s ε_i are not exactly normal. We consider two cases: (a) ε_i are uniform r.v.'s on $[0,1]$ and (b) ε_i are Laplace r.v.'s with mean zero and scale parameter 1. In both cases, the boundary crossing probability approximations developed in Chapter 2 for the probability $F_\xi(M, h) := P_0(\xi_{n,L} < h \text{ for all } n \in [0, M])$ were shown to be quite robust. This suggests Approximation 0 will also provide fairly robust approximations for the ARL in the non-normal case. Here, we shall only focus on the performance of Approximations 1-4 for the power of the test.

Simulation results are shown in Figures 5.17 and 5.18. The style of Figures 5.17 and 5.18 is very similar to many previous figures starting with Figure 5.6. The thick black dashed line corresponds to Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ computed with 100,000 simulations. The red solid lines correspond to the approximations developed in this chapter. The figures below show that the approximation remain rather accurate, at least for the chosen distributions of the r.v. ε_i .

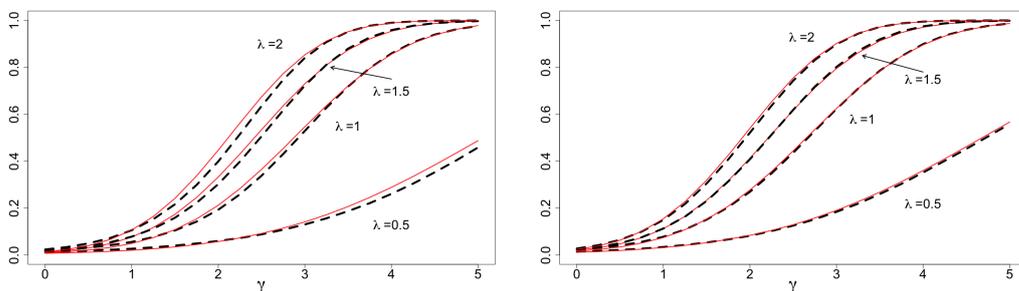


Figure 5.17: Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and its approximations (solid red). Left: $\varepsilon_i \sim \text{Laplace}(0, 1)$ with $L = 5$. Right: $\varepsilon_i \sim \text{Laplace}(0, 1)$ with $L = 20$.

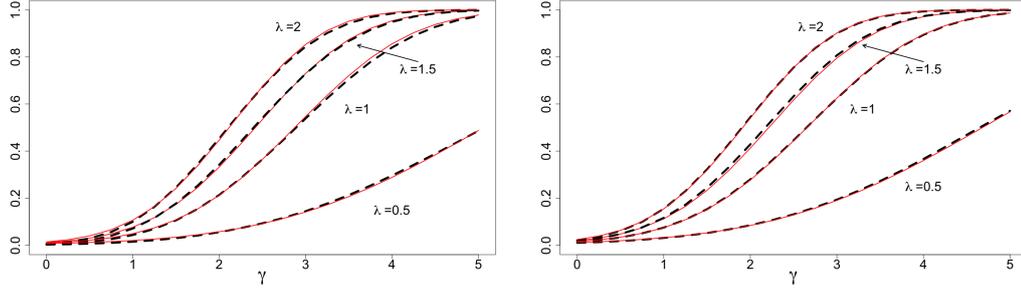


Figure 5.18: Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ (dashed black) and its (solid red). Left: $\varepsilon_i \sim \text{Uniform}[0, 1]$ with $L = 5$. Right: $\varepsilon_i \sim \text{Uniform}[0, 1]$ with $L = 20$.

5.6.2 Robustness to non-uniform weights

The author explored the following two ways of incorporating weights.

- (i) L random weights w_1, w_2, \dots, w_L , with w_i i.i.d. uniform on $[0.5, 1.5]$, are associated with a position in the moving window; this results in the moving weighted sum

$$S_{n,w,L} := \sum_{j=n+1}^{n+L} w_{j-n} \varepsilon_j;$$

- (ii) Random weights $w_1, w_2 \dots$ are associated with r.v. $\varepsilon_1, \varepsilon_2 \dots$; here w_j are i.i.d. uniform r.v.'s on $[0, 2]$; this gives the moving weighted sum

$$S_{n,w,L} := \sum_{j=n+1}^{n+L} w_j \varepsilon_j$$

Simulations results are shown in Fig. 5.19. In both cases, simulations have been repeated 500 times and plotted all the curves representing the power $\mathcal{P}_\xi(h, A, L)$ in grey colour. For simplicity, only considered the case of $\lambda = 1$ has been considered; Approximation 1 is plotted as red solid line.

5.7 Appendix A

In this section, we shall prove result (5.4.13). Using (5.4.2) we rewrite $F_{h,0,-\gamma,0,\gamma}(4|x)$ as

$$\begin{aligned} & F_{h,0,-\gamma,0,\gamma}(4|x) \\ &= \Pr\{W(t) - W(t+1) < B(t; h, 0, -\gamma, 0, \gamma) \text{ for all } t \in [0, 4] \mid W(0) - W(1) = x\} \\ &= \Pr\{W(t) - W(t+1) < h, W(t+1) - W(t+2) < h - \gamma t, W(t+2) - W(t+3) \\ &\quad < h - \gamma, W(t+3) - W(t+4) < h - \gamma + \gamma t \text{ for all } t \in [0, 1] \mid W(0) - W(1) = x\} \\ &= \Pr\{W(t) < W(t+1) + h < W(t+2) + 2h - \gamma t < W(t+3) + 3h - \gamma - \gamma t < \\ &\quad W(t+4) + 4h - 2\gamma \text{ for all } t \in [0, 1] \mid W(0) - W(1) = x\}. \end{aligned}$$

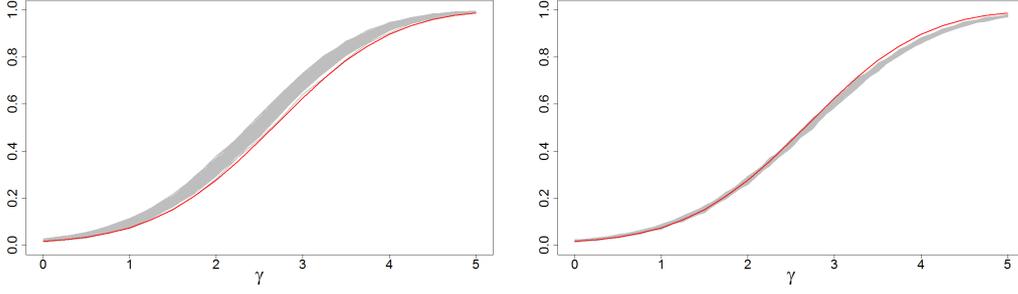


Figure 5.19: Monte Carlo approximations of $\mathcal{P}_\xi(h, A, L)$ with weights (solid grey) and Approximation 1 (solid red). Left: case (i) with $L = 20$ and $h = 3$. Right: case (ii) with $L = 20$ and $h = 3$.

Let Ω be the event defined as follows

$$\Omega = \{W(t) < W(t+1) + h < W(t+2) + 2h - \gamma t < W(t+3) + 3h - \gamma - \gamma t < W(t+4) + 4h - 2\gamma \text{ for all } t \in [0, 1]\}$$

and let $x_i = W(i)$, $i = 0, 1, \dots, 5$. Integrating out over the values x_i , by the law of total probability we obtain:

$$F_{h,0,-\gamma,0,\gamma}(4|x) = \int \dots \int \Pr\{\Omega \mid W(0) = x_0, \dots, W(5) = x_5, W(0) - W(1) = x\} \\ \times \Pr\{W(0) \in dx_0, \dots, W(5) \in dx_5 \mid W(0) - W(1) = x\}. \quad (5.7.1)$$

Note that $W(1) = x_1 = -x$, since $W(0) - W(1) = x$ and $W(0) = 0$. Define individually the following processes:

$$\begin{aligned} W_0(t) &= W(t) \\ W_1(t) &= h + W(t+1) \\ W_2(t) &= 2h - \gamma t + W(t+2) \\ W_3(t) &= 3h - \gamma - \gamma t + W(t+3) \\ W_4(t) &= 4h - 2\gamma + W(t+4) \end{aligned}$$

with $0 \leq t \leq 1$ for all processes. The event Ω can be re-written as

$$\Omega = \{W_0(t) < W_1(t) < W_2(t) < W_3(t) < W_4(t) \text{ for all } t \in [0, 1]\}.$$

The conditioning introduced in (5.7.1) results in:

$$\begin{aligned} W_0(0) &= 0 & W_0(1) &= x_1 \\ W_1(0) &= h + x_1 & W_1(1) &= h + x_2 \\ W_2(0) &= 2h + x_2 & W_2(1) &= 2h - \gamma + x_3 \\ W_3(0) &= 3h - \gamma + x_3 & W_3(1) &= 3h - 2\gamma + x_4 \\ W_4(0) &= 4h - 2\gamma + x_4 & W_4(1) &= 4h - 2\gamma + x_5. \end{aligned}$$

From this, we can express (5.7.1) as

$$\begin{aligned}
F_{h,0,-\gamma,0,\gamma}(4|x) &= \int \cdots \int \Pr\{\Omega \mid W_0(0) = 0, \dots, W_4(0) = 4h - 2\gamma + x_4, W_0(1) = x_1, \dots, \\
&\quad W_4(1) = 4h - 2\gamma + x_5, W_0(0) - W_0(1) = x\} \\
&\quad \times \Pr\{W(0) \in dx_0, \dots, W(5) \in dx_5 \mid W(0) - W(1) = x\}. \quad (5.7.2)
\end{aligned}$$

The region of integration for (5.7.2) is determined from the following inequalities:

$$x_1 < x_2 + h < x_3 + 2h - \gamma < x_4 + 3h - 2\gamma < x_5 + 4h - 2\gamma.$$

Thus, the upper limit of integration is infinity for all x_i . For integration with respect to x_5 , the lower limit is $x_4 - h$. For integration with respect x_4 , the lower limit is $x_3 - h + \gamma$. For x_3 , the lower limit is $x_2 - h + \gamma$. Finally, for x_2 the lower limit of integration in $x_1 - h = -x - h$. Now using Corollary 4.2.3 in Chapter 4 with $n = 4$ we obtain:

$$\begin{aligned}
&\Pr\{\Omega \mid W_0(0) = 0, \dots, W_4(0) = 4h - 2\gamma + x_4, \\
&\quad W_0(1) = x_1, \dots, W_4(1) = 4h - 2\gamma + x_5, W_0(0) - W_0(1) = x\} \\
&= \exp(-|\boldsymbol{\mu}|^2/2 + \boldsymbol{\mu} \cdot (\mathbf{c} - \mathbf{a})) \det[\varphi(a_i, c_j)]_{i,j=0}^4 / \prod_{i=0}^4 \varphi(a_i - c_i + \mu_i),
\end{aligned}$$

$\boldsymbol{\mu}$, \mathbf{a} and \mathbf{c} are given by:

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \\ -\gamma \\ -\gamma \\ 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ x_1 + h \\ x_2 + 2h \\ x_3 + 3h - \gamma \\ x_4 + 4h - 2\gamma \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} x_1 \\ x_2 + h \\ x_3 + 2h - \gamma \\ x_4 + 3h - 2\gamma \\ x_5 + 4h - 2\gamma \end{bmatrix}$$

The second probability in the right-hand side of (5.7.2) is $\prod_{i=1}^4 \varphi(x_i - x_{i+1})$. Using the fact

$$\prod_{i=0}^4 \varphi(a_i - c_i + \mu_i) = \prod_{i=0}^4 \varphi(x_i - x_{i+1}),$$

and collecting all results we obtain:

$$\begin{aligned}
F_{h,0,-\gamma,0,\gamma}(4|x) &= \frac{e^{\gamma^2}}{\varphi(x)} \int_{-x-h}^{\infty} dx_2 \int_{x_2-h+\gamma}^{\infty} dx_3 \int_{x_3-h+\gamma}^{\infty} dx_4 \int_{x_4-h}^{\infty} e^{-\gamma(x_4-x_2)} dx_5 \times \\
\det &\begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h+\gamma) & \varphi(-x_4-3h+2\gamma) & \varphi(-x_5-4h+2\gamma) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h+\gamma) & \varphi(-x-x_4-2h+2\gamma) & \varphi(-x-x_5-3h+2\gamma) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3+\gamma) & \varphi(x_2-x_4-h+2\gamma) & \varphi(x_2-x_5-2h+2\gamma) \\ \varphi(x_3+3h-\gamma+x) & \varphi(x_3+2h-\gamma-x_2) & \varphi(h) & \varphi(x_3-x_4+\gamma) & \varphi(x_3-x_5-h+\gamma) \\ \varphi(x_4+4h-2\gamma+x) & \varphi(x_4+3h-2\gamma-x_2) & \varphi(x_4+2h-\gamma-x_3) & \varphi(h) & \varphi(x_4-x_5) \end{bmatrix} \\
&= \frac{e^{\gamma^2}}{\varphi(x)} \int_{-x-h}^{\infty} dx_2 \int_{x_2-h+\gamma}^{\infty} dx_3 \int_{x_3-h+\gamma}^{\infty} e^{-\gamma(x_4-x_2)} dx_4 \times \\
\det &\begin{bmatrix} \varphi(x) & \varphi(-x_2-h) & \varphi(-x_3-2h+\gamma) & \varphi(-x_4-3h+2\gamma) & \Phi(-x_4-3h+2\gamma) \\ \varphi(h) & \varphi(-x-x_2) & \varphi(-x-x_3-h+\gamma) & \varphi(-x-x_4-2h+2\gamma) & \Phi(-x-x_4-2h+2\gamma) \\ \varphi(x_2+2h+x) & \varphi(h) & \varphi(x_2-x_3+\gamma) & \varphi(x_2-x_4-h+2\gamma) & \Phi(x_2-x_4-h+2\gamma) \\ \varphi(x_3+3h-\gamma+x) & \varphi(x_3+2h-\gamma-x_2) & \varphi(h) & \varphi(x_3-x_4+\gamma) & \Phi(x_3-x_4+\gamma) \\ \varphi(x_4+4h-2\gamma+x) & \varphi(x_4+3h-2\gamma-x_2) & \varphi(x_4+2h-\gamma-x_3) & \varphi(h) & \Phi(h) \end{bmatrix}.
\end{aligned}$$

□

Chapter 6

Application to the Singular Spectrum Analysis (SSA) change-point detection algorithm

Abstract

In this chapter, we study approximations of boundary crossing probabilities for the maximum of moving weighted sums of i.i.d. random variables. We consider a particular case of weights obtained from a trapezoidal weight function which, under certain parameter choices, can also result in an unweighted sum. It is demonstrated that the approximations based on classical results of extreme value theory provide some scope for improvement, particularly for a range of values required in practical applications. The problem considered in this chapter is a generalisation of the problem considered in previous chapters, in particular the boundary crossing probabilities studied in Chapter 2. The content of this chapter has been published in [80].

6.1 Introduction: Statement of the problem

This chapter considers the approximation of a more complex boundary crossing probability than the one studied in previous chapters; in particular the boundary crossing probability defined in (2.1.2) of Chapter 2. The main difference in this chapter to previous chapters, is that we consider a moving weighted sum process rather than a simple moving sum process. The particular weights in the moving sum arise from the singular spectrum analysis (SSA) change-point detection algorithm described in [74]; more details about this connection with SSA are provided in Section 6.2.2. For $Q = 1$ in $w_{L,Q}(t)$ defined later in Section 6.1.2, if the random variables under consideration are i.i.d. normal, then the moving weighted sum considered in this chapter becomes the MOSUM change-point procedure studied comprehensively in previous chapters. Therefore, the problem considered here is a generalisation but is significantly more complex. Let us formulate the problem.

Let $\varepsilon_1, \varepsilon_2, \dots$ be a sequence of independent identically distributed random variables with finite mean μ and variance σ^2 and some c.d.f. F . Define the moving

weighted sum as

$$\mathbb{S}_{n;L,Q} = \sum_{s=n+1}^{n+L+Q-1} w_{L,Q}(s-n)\varepsilon_s \quad (n = 0, 1, \dots), \quad (6.1.1)$$

where the weight function $w_{L,Q}(\cdot)$ is defined by

$$w_{L,Q}(t) = \begin{cases} t & \text{for } 0 \leq t \leq Q, \\ Q & \text{for } Q \leq t \leq L, \\ L+Q-t & \text{for } L \leq t \leq L+Q-1. \end{cases} \quad (6.1.2)$$

where L and Q are positive integers with $Q \leq L$.

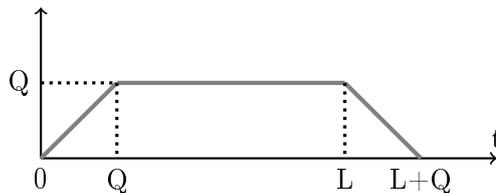


Figure 6.1: The weight function $w_{L,Q}(\cdot)$, $1 \leq Q \leq L$.

The weight function $w_{L,Q}(\cdot)$ is depicted in Figure 6.1. In the special case $Q = 1$, the weighted moving sum (6.1.1) becomes an ordinary moving sum.

The main aim of this chapter is to study precision of different approximations of boundary crossing probabilities for the maximum of the moving weighted sum; that is,

$$P \left(\max_{n=0,1,\dots,M} \mathbb{S}_{n;L,Q} > H \right), \quad (6.1.3)$$

where H is a given threshold, M is reasonably large and L, Q are fixed parameters.

This chapter is structured as follows. In Section 6.2 we reformulate the problem and provide motivation why a trapezoidal weight function is considered. In Section 6.3, a number of approximations to (6.1.3) are introduced based on the classical extreme value theory. Using the classical approximations, which do not perform very well, we also derive another approximation (called ‘combined’) which appears to be more accurate. The performance of these approximations is analyzed by a large simulation study described in Section 6.4.

6.2 Boundary crossing probabilities: discrete and continuous time

6.2.1 Reformulation of the problem

For convenience of dealing with the probability (6.1.3), we standardise the moving weighted sum $\mathbb{S}_{n;L,Q}$. Derivation of the following lemma is straightforward.

Lemma 6.2.1 *The first two moments of $\mathbb{S}_{n;L,Q}$ are*

$$E\mathbb{S}_{n;L,Q} = \mu LQ, \quad \text{var}(\mathbb{S}_{n;L,Q}) = \frac{\sigma^2 Q}{3}(3LQ - Q^2 + 1). \quad (6.2.1)$$

We now define the standardized random variables (r.v.)

$$\zeta_n := \frac{\mathbb{S}_{n;L,Q} - E\mathbb{S}_{n;L,Q}}{\sqrt{\text{var}(\mathbb{S}_{n;L,Q})}} = \frac{\sqrt{3}(\mathbb{S}_{n;L,Q} - \mu LQ)}{\sigma\sqrt{Q(3LQ - Q^2 + 1)}}, \quad (6.2.2)$$

$n = 0, 1, \dots$. If the r.v. $\varepsilon_1, \varepsilon_2, \dots$ are normal then the r.v. ζ_1, ζ_2, \dots are also normal. Otherwise, using the Central Limit Theorem, we obtain that $\zeta_n \sim N(0, 1)$ holds asymptotically, as $L \rightarrow \infty$.

Using the notation ζ_n , the problem (6.1.3) is equivalent to studying approximations for the boundary crossing probability (abbreviated BCP)

$$P_{M,h}(\zeta_n) := P\left(\max_{n=0,1,\dots,M} \zeta_n > h\right), \quad (6.2.3)$$

where

$$H = \mu LQ + \sigma h \sqrt{\frac{Q(3LQ - Q^2 + 1)}{3}}.$$

A number of approaches could be used to approximate (6.2.3). We could have ignored the dependence structure of the sequence of moving weighted sums and used either asymptotic normality alone or the limiting extreme value distribution to choose h . Instead, in what follows we study several approximations of (6.2.3) which are based on approximating the sequence ζ_n by a continuous time random process. Before we proceed, let us consider a special case of ε_j , which has important practical significance.

6.2.2 Motivation for the problem

If we let $\varepsilon_j = \xi_j^2$, where ξ_1, ξ_2, \dots are i.i.d. random variables with zero mean, variance δ^2 and finite fourth moment $\mu_4 = E\xi_i^4$, then $\mathbb{S}_{n;L,Q}$ can be seen as a moving weighted sum of squares. In this case, the mean $\mu = E\varepsilon_j = \delta^2$ and $\sigma^2 = \text{var}(\varepsilon_j) = \mu_4 - \delta^4$. By approximating (6.1.3) we are considering a particularly interesting case linked to the SSA change-point detection algorithm proposed in [74]. A good approximation for the BCP for the maximum of the moving weighted sums of squares is needed in the theory of sequential change-point detection because the BCP defines the significance levels for the SSA change-point detection statistic. For an extensive introduction to SSA, see [35, 36].

6.2.3 Continuous time approximation

By the definition, the probability $P_{M,h}(\zeta_n)$ is an $(M+1)$ -dimensional integral which is difficult to compute. We assume that $L \rightarrow \infty$ and consider a transformation described below in Section 6.3 from the time series ζ_n , $n = 0, 1, \dots, M$, to a continuous-time process ζ_t , $t \in [0, T]$, where $T = M/\sqrt{LQ}$ for large Q , see (6.3.2), and $T = M/L$

in the case of small Q , see beginning of Section 6.3.2. Like the time series ζ_n , the process ζ_t is standardized so that $E\zeta_t = 0$ and $E\zeta_t^2 = 1$ for all t . Also, the process ζ_t is Gaussian and stationary with some autocorrelation function $R(s) = E\zeta_0\zeta_s$.

By such a transformation, the probability $P_{M,h}(\zeta_n)$ is approximated by $P(T, h, \zeta_t)$, which is the probability of reaching the threshold h by the process ζ_t on the interval $[0, T]$; that is,

$$\begin{aligned} P_{M,h}(\zeta_n) &\cong P(T, h, \zeta_t) = \Pr \left\{ \max_{0 \leq t \leq T} \zeta_t \geq h \right\} \\ &= \Pr \left\{ \zeta_t \geq h \text{ for at least one } t \in [0, T] \right\}. \end{aligned} \quad (6.2.4)$$

For the continuous process ζ_t , two main useful characteristics are the probability density function of reaching the threshold h for the first time

$$q(t, h, \zeta_t) = \frac{d}{dt}P(t, h, \zeta_t), \quad 0 < t < \infty, \quad (6.2.5)$$

and the average time $\varrho(h, \zeta_t)$ until the process ζ_t reaches the threshold h

$$E(\varrho(h, \zeta_t)) = \int_0^\infty tq(t, h, \zeta_t)dt = \int_0^\infty tdP(t, h, \zeta_t).$$

From the practical point of view, we are interested in finding good approximations of (6.2.3) for small and moderate M . But the mathematical theory guarantees accurate approximations just for large M .

To proceed further, we need to discuss results concerning the autocorrelation function of the continuous process ζ_t . This can be done through computing the correlations between $\mathbb{S}_{n;L,Q}$ and $\mathbb{S}_{n+\nu;L,Q}$ for $\nu > 0$.

6.2.4 Correlation between $\mathbb{S}_{n;L,Q}$ and $\mathbb{S}_{n+1;L,Q}$

For fixed L and Q , the moving weighted sum $\mathbb{S}_{n;L,Q}$ is a function of n . The index n can be treated as time and thus the sequence $\mathbb{S}_{0;L,Q}, \mathbb{S}_{1;L,Q}, \dots$ defined in (6.1.1) can be considered as a time series. In order to derive the approximations of this chapter, we need explicit expressions for the correlation $\text{Corr}(\mathbb{S}_{n;L,Q}, \mathbb{S}_{n+1;L,Q})$. The general case $\text{Corr}(\mathbb{S}_{n;L,Q}, \mathbb{S}_{n+\nu;L,Q})$, $\nu > 1$ need not be considered for these approximations.

Without loss of generality, we can assume that $n = 0$ and we denote $\mathbb{S}_\nu := \mathbb{S}_{\nu;L,Q}$ where $\nu = 0, 1$.

Lemma 6.2.2 *The correlation $\text{Corr}(\mathbb{S}_0, \mathbb{S}_1) = \text{Corr}(\mathbb{S}_{n;L,Q}, \mathbb{S}_{n+1;L,Q})$, where $\mathbb{S}_{n;L,Q}$ is defined in (6.1.1), is*

$$\text{Corr}(\mathbb{S}_0, \mathbb{S}_1) = \frac{E(\mathbb{S}_0\mathbb{S}_1) - (E\mathbb{S}_0)^2}{\text{var}(\mathbb{S}_0)} = 1 - \frac{3}{3LQ - Q^2 + 1}.$$

Proof. From the definition (6.1.1), the quadratic forms \mathbb{S}_0 and \mathbb{S}_1 can be represented as

$$\mathbb{S}_0 = \sum_{i=1}^{Q-1} i\varepsilon_i + Q \sum_{i=Q}^L \varepsilon_i + \sum_{i=L+1}^{Q+L-1} (Q+L-i)\varepsilon_i$$

and

$$\mathbb{S}_1 = \mathbb{S}_0 - \sum_{i=1}^Q \varepsilon_i + \sum_{i=1}^Q \varepsilon_{L+i}.$$

Using these representations, we can easily obtain $E(\mathbb{S}_0\mathbb{S}_1) = E\mathbb{S}_0^2 - Q\sigma^2$. Then by substituting the explicit expressions (6.2.1) for $E\mathbb{S}_0$ and $\text{var}(\mathbb{S}_0) = E\mathbb{S}_0^2$, we obtain the desired result. \square

Note that the correlation does not depend on the distribution of errors ε_j (unlike the covariance which depends on the mean μ and variance σ^2 of ε_j). This also can be seen in relation to the fact (see, for example, [104]) that the spectral density of the moving average process depends only on the weight function, which is $w_{L,Q}(t)$ in our case.

6.3 Approximations of the boundary crossing probabilities

In this section we formulate four different approximations for the BCP $P_{M,h}(\zeta_n)$ defined in (6.2.4). These approximations depend on the behaviour of the autocorrelation function $R(s) = E\zeta_0\zeta_s$ at 0 which in its turn depends on parameters Q and L of the weight function in (6.1.2). We consider the following two cases: (i) large Q and large L , (ii) small Q and large L .

6.3.1 Case of large Q and large L

Consider the sequence of random variables $\zeta_0, \zeta_1, \dots, \zeta_M$ defined in (6.2.2). In view of Lemma 6.2.2, the correlation between ζ_n and ζ_{n+1} is

$$\text{Corr}(\zeta_n, \zeta_{n+1}) = 1 - \frac{3}{3LQ - Q^2 + 1}. \quad (6.3.1)$$

Assume that both L and Q are large. Moreover, assume that L and Q tend to infinity in such a way that the limit $\lambda = \lim Q/L$ exists and $0 < \lambda \leq 1$. Set $\Delta = 1/\sqrt{LQ}$ and

$$t_n = n\Delta, \quad n = 0, 1, \dots, M, \quad \text{so that } t_n \in [0, T] \text{ with } T = M\Delta. \quad (6.3.2)$$

Define a piece-wise linear continuous-time process $\zeta_t^{(L)}$, $t \in [0, T]$, as follows

$$\zeta_t^{(L)} = \frac{1}{\Delta} [(t_n - t)\zeta_{n-1} + (t - t_{n-1})\zeta_n] \quad \text{for } t \in [t_{n-1}, t_n], \quad n = 1, \dots, M. \quad (6.3.3)$$

By construction, the process $\zeta_t^{(L)}$ is such that $\zeta_{t_n}^{(L)} = \zeta_n$ for $n = 0, \dots, M$. Also we have that $\zeta_t^{(L)}$ is a second-order stationary process in the sense that $E\zeta_t^{(L)}$, $\text{var}(\zeta_t^{(L)})$ and the autocorrelation function $R_\zeta^{(L)}(t, t + k\Delta) = \text{Corr}(\zeta_t^{(L)}, \zeta_{t+k\Delta}^{(L)})$ do not depend on t .

Lemma 6.3.1 *Let $\lambda = \lim_{L,Q \rightarrow \infty} Q/L$ and assume that $0 < \lambda \leq 1$. Consider the process $\zeta_t^{(L)}$ defined in (6.3.3). The limiting process $\zeta_t = \lim_{L,Q \rightarrow \infty} \zeta_t^{(L)}$ is stationary Gaussian with some autocorrelation function $R_\zeta(t, t + s) = R(s)$. Moreover, $R'(0) = 0$ and $R''(0) = -6/(3 - \lambda)$.*

Proof. For the autocorrelation function $R(\cdot)$ we have $R'(0) = 0$ since

$$R'(0-) = R'(0+) = \lim_{L, Q \rightarrow \infty} \frac{R(\Delta) - 1}{\Delta} = \lim_{L, Q \rightarrow \infty} \frac{-3\sqrt{LQ}}{3LQ - Q^2 + 1} = 0,$$

where we used the relations $\Delta = 1/\sqrt{LQ}$, $R(\Delta) = 1 - 3/(3LQ - Q^2 + 1)$ and $R(0) = 1$. We similarly obtain

$$R''(0) = \lim_{L, Q \rightarrow \infty} \frac{R(\Delta) + R(-\Delta) - 2R(0)}{\Delta^2} = \lim_{L, Q \rightarrow \infty} \frac{-6LQ}{3LQ - Q^2 + 1} = -\frac{6}{3-\lambda} < 0.$$

□

For a Gaussian stationary process ζ_t with $E\zeta_t = 0$ and $E\zeta_t^2 = 1$ and autocorrelation function $R(\cdot)$ such that $R'(0) = 0$ and $R''(0) < 0$ we can use the following two well-known approximations.

Approximation 1 (App 1). From Theorem 8.2.7 in [63] we have

$$\lim_{T \rightarrow \infty} P \left\{ \max_{0 \leq t \leq T} \zeta_t \leq \underbrace{\frac{u + \log \frac{\sqrt{-R''(0)}}{2\pi}}{\sqrt{2 \log T}} + \sqrt{2 \log T}}_h \right\} = \exp(-e^{-u}).$$

Expressing u in terms of h , we obtain the approximation 1

$$P(T, h, \zeta_t) \cong 1 - \exp(-e^{-u}) \quad (6.3.4)$$

with $u = \gamma(h - \gamma) + c$, where

$$\gamma = \sqrt{2 \log T} \quad \text{and} \quad c = -\log \frac{\sqrt{-R''(0)}}{2\pi} = -\log \frac{1}{2\pi} \sqrt{\frac{6}{3-\lambda}}. \quad (6.3.5)$$

Approximation 2 (App 2). From [20], we have

$$\lim_{T \rightarrow \infty} P \left\{ \max_{0 \leq t \leq T} \zeta_t \leq \underbrace{\sqrt{2 \log \mu} + \frac{v}{\sqrt{2 \log \mu}}}_h \right\} = \exp(-e^{-v}),$$

where

$$\mu = \frac{T \sqrt{-R''(0)}}{2\pi} = \frac{T}{2\pi} \sqrt{\frac{6}{3-\lambda}}.$$

Expressing v in terms of h , we obtain the approximation 2

$$P(T, h, \zeta_t) \cong 1 - \exp(-e^{-v}) \quad (6.3.6)$$

with

$$v = \sqrt{2 \log \mu} (h - \sqrt{2 \log \mu}).$$

Note that $2 \log \mu = \gamma^2 - 2c$ and

$$\sqrt{2 \log \mu} = \sqrt{\gamma^2 - 2c} = \gamma - \frac{c}{\gamma} + O\left(\frac{1}{\gamma^3}\right),$$

as $\gamma \rightarrow \infty$, where γ and c are defined in (6.3.5). Therefore, for large T (and, therefore, large γ) we have

$$v \cong \left(\gamma - \frac{c}{\gamma}\right) \left(h - \gamma + \frac{c}{\gamma}\right) = \underbrace{(h - \gamma)\gamma + c}_u - \frac{(h - \gamma)c}{\gamma} - \frac{c^2}{\gamma^2}.$$

Let us construct another approximation by combining the approximations 1 and 2.

Approximation 3 (combined). Consider the approximation

$$P(T, h, \zeta_t) \cong 1 - \exp(-e^{-z}) \quad (6.3.7)$$

where

$$z = \begin{cases} u - \frac{(h-\gamma)c}{\gamma} - \frac{c^2}{\gamma^2} & \text{for } h \leq \gamma - \frac{c}{\gamma}, \\ u & \text{for } h \geq \gamma - \frac{c}{\gamma}. \end{cases}$$

Formally, $\lambda = \lim_{L, Q \rightarrow \infty} Q/L = 0$ still satisfies Lemma 6.3.1 in the sense that $R'(0) = 0$ and $R''(0) = -2 < 0$; however, the above approximations are poor when Q is small; this shall be demonstrated in Section 6.4. The case of small Q should be treated differently and is considered in the following subsection.

6.3.2 Case of small Q and large L

Consider again the sequence of random variables ζ_n defined by (6.2.2). Unlike in Section 6.3.1, now we look at the asymptotic transformation when $L \rightarrow \infty$ but Q is fixed. Set $\Delta = 1/L$ and $T = M\Delta$. Define t_n , $n = 0, 1, \dots, M$, as in (6.3.2) and consider the piece-wise linear continuous-time process $\zeta_t^{(L)}$ defined by (6.3.3).

Lemma 6.3.2 *Let Q be fixed. The limiting process ζ_t as $L \rightarrow \infty$ is a Gaussian second-order stationary process with autocorrelation function $R_\zeta(t, t+s) = R(s)$. Moreover, $R'(0+) = -\frac{1}{Q} \neq 0$.*

Proof. We first note that

$$\left. \frac{\partial R_\zeta(t, s)}{\partial s} \right|_{s=t+} = R(0+).$$

Using (6.3.1) and the fact that $\Delta = 1/L$, we have

$$R'(0+) = \lim_{L \rightarrow \infty} \frac{R(\Delta) - R(0)}{\Delta} = - \lim_{L \rightarrow \infty} \frac{3L}{3LQ - Q^2 + 1} = -\frac{1}{Q}.$$

□

Let us now formulate the tangent approximation suggested in [22]; it is one of the most known approximations for the density function $q(t, h, \zeta_t)$ of the first passage time defined in (6.2.5). Using this, we can approximate the first passage probability $P(T, h, \zeta_t)$ defined in (6.2.4) in the case of a Gaussian process $\zeta(t)$ on $[0, T]$ with $E\zeta(t)=0$, some autocorrelation function $R_\zeta(t, s)$ and the possibly non-constant

threshold $h = h(t)$.

The Durbin approximation for $q(t, h, \zeta_t)$ can be written as

$$q(t, h, \zeta_t) \cong b_0(t, h)f(t, h),$$

where

$$f(t, h) = \frac{1}{\sqrt{2\pi R_\zeta(t, t)}} e^{-\frac{h^2(t)}{2R_\zeta(t, t)}}, \quad b_0(t, h) = -\frac{h(t)}{R_\zeta(t, t)} \frac{\partial R_\zeta(s, t)}{\partial s} \Big|_{s=t+} - \frac{dh(t)}{dt}.$$

In view of (6.2.5) the related approximation for the first passage probability $P(T, h, \zeta_t)$ is

$$P(T, h, \zeta_t) \cong \int_0^T b_0(t, h)f(t, h)dt.$$

In the case when the threshold $h(t) = h$ is constant, using Lemma 6.3.2 we obtain

$$b_0(t, h) = -hR'(0+) = \frac{h}{Q}, \quad q(t, h, \zeta_t) \cong \frac{h}{\sqrt{2\pi Q}} e^{-h^2/2}$$

and therefore we obtain the following approximation.

Approximation 4 (App 4). The Durbin approximation for the BCP (6.2.4) is

$$P(T, h, \zeta_t) \cong \frac{hT}{\sqrt{2\pi Q}} e^{-h^2/2}. \quad (6.3.8)$$

6.4 Simulation study

In this section we study quality of approximations for the BCP $P_{M,h}(\zeta_n)$ defined in (6.2.3), where ε_t are normal r.v.'s with mean 0 and variance 1. Asymptotically (for large L and M), the approximations we study can also be used for the BCP connected to the weighted sum of squares discussed in Section 2.2 and therefore for setting significance levels for the SSA change-point statistic defined in [74].

In Figures 6.2–6.6, the 'Sum of normal' line corresponds to the empirical value of (6.2.3) computed from 100,000 simulations with different values of L , Q and M . In simulations leading to Figures 6.2–6.4 the value of Q can be considered as large and hence we compare Approximations 1–3. In Figure 6.5 we present analysis demonstrating the lack of accuracy of Approximations 1–3 when Q is small. We then analyse the performance of the Durbin approximation in Figure 6.6, which is constructed specifically under the assumption that Q is small; in this case we set $Q = 1$. We observe that for large L and Q approximation 3 is typically superior to the approximations 1 and 2 for all h (note that approximations 1 and 3 coincide for large values of h). Listed in Tables 6.1–6.4 are the approximated threshold values h (for approximations 1 and 2 only) for a specified true BCP, when this BCP is small enough. In these tables, R.E. denotes the relative error.

As seen in Figure 6.2, for the chosen parameters approximation 2 is generally poor; for small BCP we see particularly high relative errors in Table 6.1. On the other hand, approximation 1 performs well for small BCP and, although discrepancies can be seen for small h , we see that approximation 3 performs quite well across all values of h .

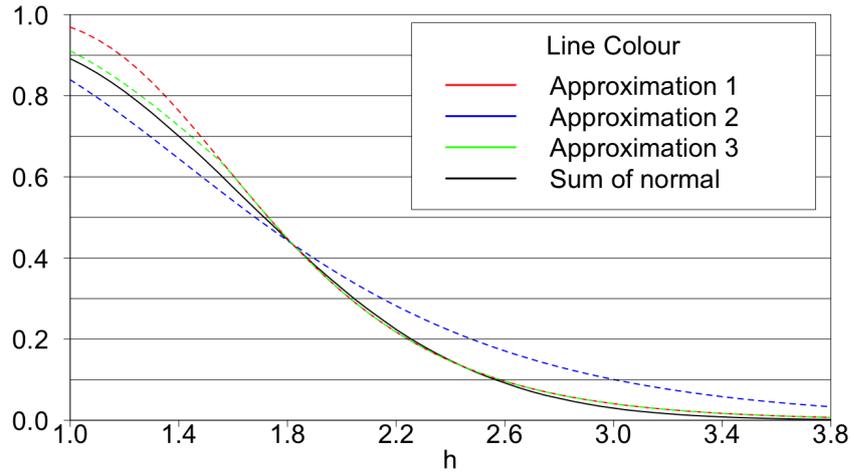


Figure 6.2: The BCP for the weighted sum of normal r.v. and its approximations: $L = 150$, $Q = 50$, $M = 1000$.

BCP	Sum of normal	App 1	App 2	R.E. for App 1 (%)	R.E. for App 2 (%)
0.05	2.833	2.907	3.510	2.612	23.897
0.10	2.572	2.582	3.004	0.389	16.796
0.15	2.401	2.386	2.700	0.625	12.453
0.20	2.264	2.243	2.477	0.928	9.408

Table 6.1: Threshold for a given BCP for the weighted sum of normal r.v. and approximations: $L = 150$, $Q = 50$, $M = 1000$

Approximation 2, whilst still being considerably worse than approximations 1 and 3, shows signs of improvement with this choice of L and Q . At the BCP of 0.05, approximation 1 produces the lowest relative error with the parameter choices considered so far.

Table 6.2: Threshold for a given BCP for the weighted sum of normal r.v. and its approximations: $L = 100$, $Q = 50$, $M = 1000$

BCP	Sum of normal	App 1	App 2	R.E. for App 1 (%)	R.E. for App 2 (%)
0.05	2.911	2.984	3.460	2.508	18.859
0.10	2.654	2.671	3.004	0.641	13.188
0.15	2.491	2.483	2.730	0.321	9.595
0.20	2.362	2.345	2.530	0.720	7.113

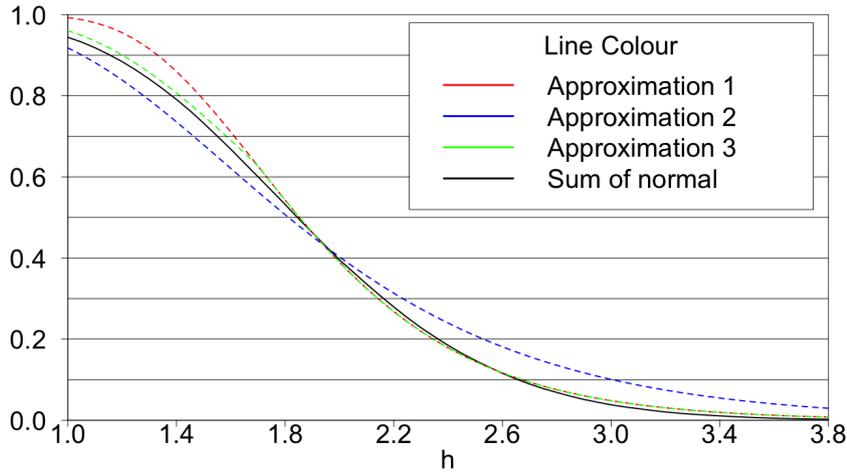


Figure 6.3: The BCP for the weighted sum of normal r.v. and approximations: $L = 100$, $Q = 50$, $M = 1000$.

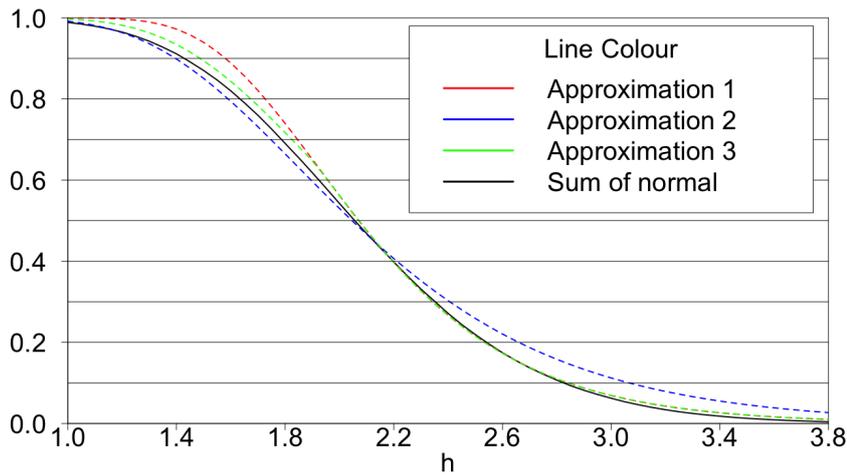


Figure 6.4: The BCP for the weighted sum of normal r.v. and its approximations: $L = 100$, $Q = 100$, $M = 2000$.

We see a considerable improvement in approximation 2 with the increase in M from 1000 to 2000, however approximation 3 still remains far superior. For this larger M , approximation 1 shows the smallest relative error at a BCP of 0.05 which is arguably the most important case.

We shall now consider the performance of approximations 1–3 for small Q . We conclude that all three approximations perform poorly when Q is not large enough (of order L).

Table 6.3: Threshold for a given BCP for the weighted sum of normal r.v. and approximations: $L = 100, Q = 100, M = 2000$

BCP	Sum of normal	App 1	App 2	R.E. for App 1 (%)	R.E. for App 2 (%)
0.05	3.063	3.135	3.455	2.351	12.798
0.10	2.816	2.841	3.066	0.888	8.878
0.15	2.659	2.664	2.831	0.188	6.469
0.20	2.541	2.534	2.660	0.275	4.683

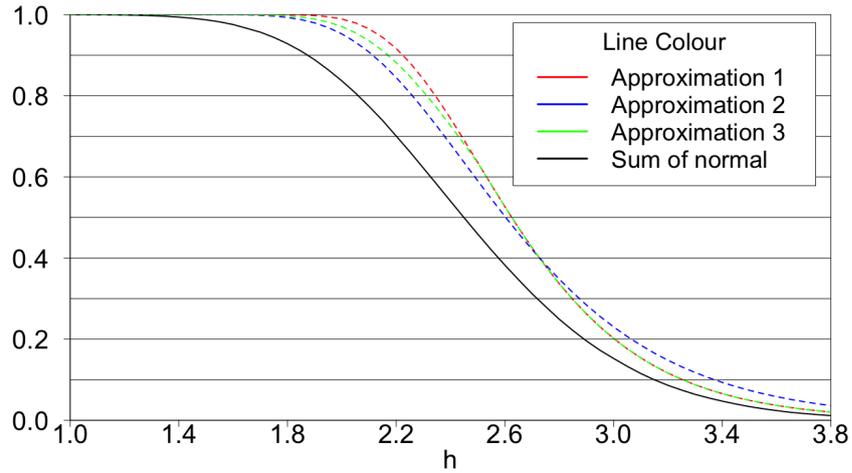


Figure 6.5: The BCP for the weighted sum of normal r.v. and its approximations: $L = 100, Q = 5, M = 2000$.

As can be seen from Figure 5 and Table 4, all three approximations are poor for $Q = 5$. Relative errors are high and thus the use of these approximations for the case of small Q and large L cannot be justified.

Table 6.4: Threshold for a given BCP for the weighted sum of normal r.v. and approximations: $L = 100, Q = 5, M = 2000$

BCP	Sum of normal	App 1	App 2	R.E. for App 1 (%)	R.E. for App 2 (%)
0.05	3.380	3.494	3.664	3.373	8.402
0.10	3.150	3.254	3.371	3.302	7.016
0.15	3.007	3.109	3.194	3.392	6.219
0.20	2.896	3.004	3.065	3.729	5.836

For checking the quality of the Durbin approximation we will use the same settings as for the approximations 1, 2 and 3. In Figure 6.6, we show results for the Durbin approximation for a few particular values of L and Q .

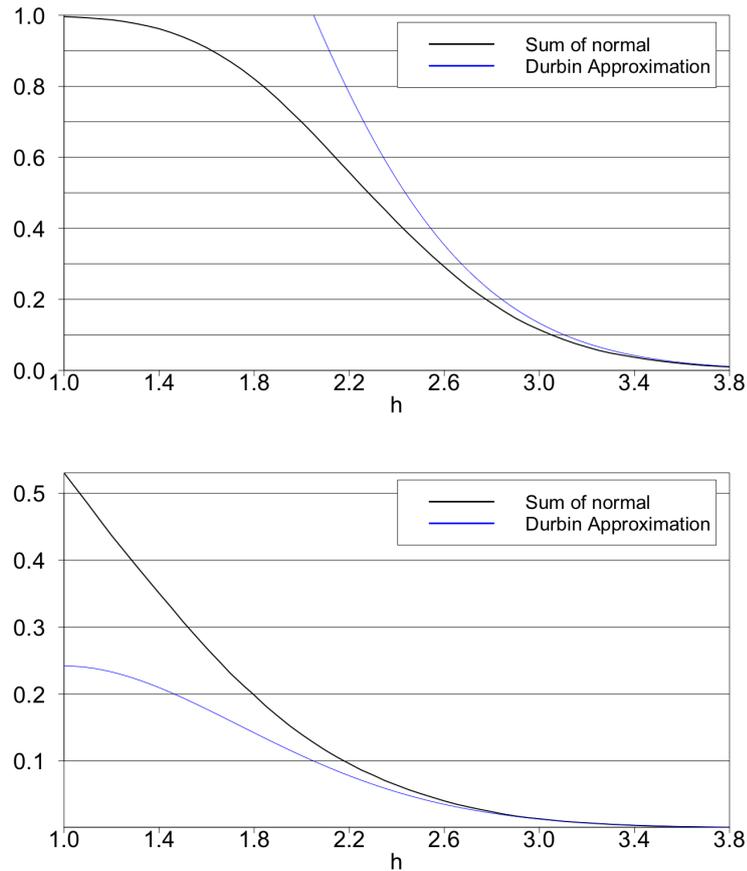


Figure 6.6: The BCP for the weighted sum of normal random variables and the Durbin approximation: $L = 300$, $Q = 1$, $T = 10$ (top) and $L = 300$, $Q = 1$, $T = 1$ (bottom)

We can conclude that the quality of the Durbin approximation (6.3.8) is poor unless the threshold h is very large. This is seen graphically in Figure 6.6 as well as numerically in Table 6.5, where there is a sharp increase in the relative error as the BCP increases. For the BCP of 0.05 the relative error for the Durbin approximation is higher than all relative errors of approximation 1 considered in this chapter.

Table 6.5: Threshold for a given BCP for the weighted sum of normal r.v. and Durbin approximation: $L = 300$, $Q = 1$, $T = 1$

BCP	Sum of normal	Durbin Approx.	R.E. for Durbin Approx. (%)
0.05	2.520	2.436	3.333
0.10	2.190	2.049	6.438
0.15	1.970	1.756	10.863
0.20	1.794	1.464	18.395

Part II

High-dimensional Covering

Chapter 7

Covering of high-dimensional cubes and quantization

Abstract

In this chapter, as the main problem we consider covering of a d -dimensional cube by n balls with reasonably large d (10 or more) and reasonably small n , like $n = 100$ or $n = 1000$. We do not require the full coverage but only 90% or 95% coverage. We establish that efficient covering schemes have several important properties which are not seen in small dimensions and in asymptotical considerations, for very large n . One of these properties can be termed ‘do not try to cover the vertices’ as the vertices of the cube and their close neighbourhoods are very hard to cover and for large d there are far too many of them. We clearly demonstrate that, contrary to a common belief, placing balls at points which form a low-discrepancy sequence in the cube, results in a very inefficient covering scheme. For a family of random coverings, we are able to provide very accurate approximations to the coverage probability. We then extend our results to the problems of coverage of a cube by smaller cubes and quantization, the latter being also referred to as facility location. Along with theoretical considerations and derivation of approximations, we provide results of a large-scale numerical investigation. The content of this chapter has been published in [144].

7.1 Introduction

In this chapter, we develop and study efficient schemes for covering and quantization in high-dimensional cubes. In particular, we will demonstrate that the proposed schemes are much superior to the so-called ‘low-discrepancy sequences’. This chapter starts by introducing the main notation. Then we formulate the main problem of covering a d -dimensional cube by n Euclidean balls. This is followed by a discussion on the main principles adopted for construction of algorithms in this chapter. Then we briefly formulate problems of covering a cube by smaller cubes (which are balls in the L_∞ -norm) and the problem of quantization. Both problems have many similarities with the main problem of covering a cube by n balls. At the end of this section, we describe the structure of the remaining sections of the chapter and summarise the main findings of the chapter.

7.1.1 Main notation

- \mathbb{R}^d : d -dimensional space; vectors in \mathbb{R}^d are row-vectors;
- $\|\cdot\|$ and $\|\cdot\|_\infty$: Euclidean and L_∞ -norms in \mathbb{R}^d ;
- $\mathcal{B}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| \leq r\}$: d -dimensional ball of radius r centered at $Z \in \mathbb{R}^d$;
- $\mathcal{B}_d(r) = \mathcal{B}_d(0, r) = \{Y \in \mathbb{R}^d : \|Y\| \leq r\}$;
- $\mathcal{S}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| = r\}$: d -dimensional sphere of radius r centered at $Z \in \mathbb{R}^d$;
- $\mathcal{C}_d(Z, \delta) = \{Y \in \mathbb{R}^d : \|Y - Z\|_\infty \leq \delta\}$: d -dimensional cube of side length 2δ centered at Z (it is also the d -dimensional ball in the L_∞ -norm with radius δ and center Z);
- $\mathcal{C}_d(\delta) = [-\delta, \delta]^d = \mathcal{C}_d(0, \delta)$;
- $\mathcal{C}_d = [-1, 1]^d = \mathcal{C}_d(1)$.

7.1.2 Main problem of interest

The main problem discussed in this chapter is the following problem of covering a cube by n balls. Let $\mathcal{C}_d = [-1, 1]^d$ be a d -dimensional cube, Z_1, \dots, Z_n be some points in \mathbb{R}^d and $\mathcal{B}_d(Z_j, r)$ be the corresponding balls of radius r centered at Z_j ($j = 1, \dots, n$). The dimension d , the number of balls n and their radius r could be arbitrary.

We are interested in the problem of choosing the locations of the centers of the balls Z_1, \dots, Z_n so that the union of the balls $\cup_j \mathcal{B}_d(Z_j, r)$ covers the largest possible proportion of the cube \mathcal{C}_d . That is, we are interested in choosing a design (a collection of points) $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ so that

$$C_d(\mathbb{Z}_n, r) := \text{vol}(\mathcal{C}_d \cap \mathcal{B}_d(\mathbb{Z}_n, r)) / 2^d \quad (7.1.1)$$

is as large as possible (given n , r and the freedom we are able to use in choosing Z_1, \dots, Z_n). Here

$$\mathcal{B}_d(\mathbb{Z}_n, r) = \bigcup_{j=1}^n \mathcal{B}_d(Z_j, r) \quad (7.1.2)$$

and $C_d(\mathbb{Z}_n, r)$ is the proportion of the cube \mathcal{C}_d covered by the balls $\mathcal{B}_d(Z_j, r)$ ($j = 1, \dots, n$). If points $Z_j \in \mathbb{Z}_n$ are random then by $C_d(\mathbb{Z}_n, r)$ we will mean $\mathbb{E}_{\mathbb{Z}_n} C_d(\mathbb{Z}_n, r)$ but we are not going to stress this in notation unless it is important.

For a design \mathbb{Z}_n , its covering radius is defined by $\text{CR}(\mathbb{Z}_n) = \max_{X \in \mathcal{C}_d} \min_{Z_j \in \mathbb{Z}_n} \|X - Z_j\|$. In computer experiments, covering radius is called minimax-distance criterion, see [47] and [105]; in the theory of low-discrepancy sequences, covering radius is called dispersion, see [79, Ch. 6].

The problem of optimal covering of a cube by n balls has very high importance for the theory of global optimization and many branches of numerical mathematics.

In particular, the n -point designs \mathbb{Z}_n with smallest CR provide the following: (a) the n -point min-max optimal quadratures, see [128, Ch.3,Th.1.1], (b) min-max n -point global optimization methods in the set of all adaptive n -point optimization strategies, see [128, Ch.4,Th.2.1], and (c) worst-case n -point multi-objective global optimization methods in the set of all adaptive n -point algorithms, see [147]. In all three cases, the class of (objective) functions is the class of Lipschitz functions, where the Lipschitz constant may be unknown. The results (a) and (b) are the celebrated results of A.G. Sukharev obtained in the late nineteen-sixties, see e.g. [127], and (c) is a recent result of A. Žilinskas, see [147].

If d is not small (say, $d > 5$) then computation of the covering radius $\text{CR}(\mathbb{Z}_n)$ for any non-trivial design \mathbb{Z}_n is a very difficult computational problem. This explains why the problem of construction of optimal n -point designs with smallest covering radius is notoriously difficult, see for example recent surveys [133, 134].

If $r = \text{CR}(\mathbb{Z}_n)$, then $C_d(\mathbb{Z}_n, r)$ defined in (7.1.1) is equal to 1, and the whole cube \mathcal{C}_d gets covered by the balls. However, we are only interested in reaching the values like 0.9, when a large part of the ball is covered. There are two main reasons why we are not interested in reaching the value $C_d(\mathbb{Z}_n, r) = 1$: (a) practical impossibility of making a numerical checking of the full coverage, if d is large enough, and (b) our approximations lose accuracy when $C_d(\mathbb{Z}_n, r)$ closely approaches 1.

If, for a given $\gamma \in [0, 1)$, we have $C_d(\mathbb{Z}_n, r) \geq 1 - \gamma$, then the corresponding coverage of \mathcal{C}_d will be called $(1-\gamma)$ -coverage; the corresponding value of r can be called $(1-\gamma)$ -covering radius. If $\gamma = 0$ then the $(1-\gamma)$ -coverage becomes the full coverage and 1-covering radius of \mathbb{Z}_n becomes $C_d(\mathbb{Z}_n, r)$. Of course, for any $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ we can reach $C_d(\mathbb{Z}_n, r) = 1$ by means of increasing r . Likewise, for any given r we can reach $C_d(\mathbb{Z}_n, r) = 1$ by sending $n \rightarrow \infty$. However, we are not interested in very large values of n and try to get the coverage of the most part of the cube \mathcal{C}_d with the radius r as small as possible. We will keep in mind the following typical values of d and n : $d = 10, 20, 50$; $n = 64, 128, 512, 1024$. Correspondingly, we will illustrate our results in such scenarios.

7.1.3 Two contradictory guiding principles and a compromise

In choosing $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$, the following two main guiding principles must be followed:

- (i) the volumes of intersections of the cube \mathcal{C}_d and each individual ball $\mathcal{B}_d(Z_j, r)$ are not very small;
- (ii) the volumes of intersections $\mathcal{B}_d(Z_j, r) \cap \mathcal{B}_d(Z_i, r)$ are small for all $i \neq j$ ($i, j = 1, \dots, n$).

These two principles do not agree with each other. Indeed, intuitively and as shown in Section 7.2 (see formulas (7.2.9)–(7.2.12)), the volume of intersection of the ball $\mathcal{B}_d(Z, r)$ and the cube \mathcal{C}_d decreases with an increasing $\|Z\|$ and hence criterion (i) favours Z_j with small norms. However, if at least some of the points Z_j get close to 0, then the distance between these points gets small and, in view of the formulas of Section 7.6.7, the volumes of intersections $\mathcal{B}_d(Z_j, r) \cap \mathcal{B}_d(Z_i, r)$ get large.

This yields that the above two guiding principles require a compromise in the rule of choosing $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ as the points Z_j should not be too far from 0 but at the same time, not too close. In particular, and this is clearly demonstrated in many examples throughout this chapter, the so-called ‘uniformly distributed sequences of points’ in \mathcal{C}_d , including ‘low-discrepancy sequences’ in \mathcal{C}_d , provide poor covering schemes. This is in a sharp contrast with the asymptotic case $n \rightarrow \infty$ (and hence $r \rightarrow 0$), when one of the recommendations, see [44, p.84], is to choose Z_j ’s from a uniformly distributed sequence of points from a set which is slightly larger than \mathcal{C}_d ; this is to facilitate covering of the boundary of \mathcal{C}_d , as it is much easier to cover the interior of the cube \mathcal{C}_d than its boundary.

In our considerations, n is not very large and hence the radius of balls r cannot be small. One recommendation for choosing $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ is to choose Z_j ’s at random in a cube $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$ (with $0 < \delta \leq 1$) with components distributed according to a suitable Beta-distribution. The optimal value of δ is always smaller than 1 and depends on d and n . If d is small or n is astronomically large, then the optimal value of δ could be close to 1 but in most interesting instances this value is significantly smaller than 1. This implies that the choice $\delta = 1$ (for example, if Z_j ’s form a uniformly distributed sequence of points in the whole cube \mathcal{C}_d) often leads to very poor covering schemes, especially when the dimension d is large (see Tables 7.1–7.3 in discussed in Section 7.3). More generally, we show that for construction of efficient designs $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$, either deterministic or randomized, we have to restrict the norms of the design points Z_j . We will call this principle ‘ δ -effect’.

7.1.4 Covering a cube by smaller cubes and quantization

In Section 7.4 we consider the problem of $(1 - \gamma)$ -coverage of the cube $\mathcal{C}_d = [-1, 1]^d$ by smaller cubes (which are L_∞ -balls). The problem of 1-covering of cube by cubes has attracted a reasonable attention in mathematical literature, see e.g. [45, 58]. The problem of $(1 - \gamma)$ -coverage of a cube by cubes happened to be simpler than the main problem of $(1 - \gamma)$ -coverage of a cube by Euclidean balls and we have managed to derive closed-form expressions for (a) the volume of intersection of two cubes, and (b) $(1 - \gamma)$ coverage, the probability of covering a random point in \mathcal{C}_d by n cubes $\mathcal{C}_d(Z_i, r)$ for a wide choice of randomized schemes of choosing designs $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$. The results of Section 7.4 show that the δ -effect holds for the problem of coverage of the cube by smaller cubes in the same degree as for the main problem of Section 7.3 of covering with balls.

Section 7.5 is devoted to the following problem of quantization also known as the problem of facility location. Let $X = (x_1, \dots, x_d)$ be uniform on $\mathcal{C}_d = [-1, 1]^d$ and $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ be an n -point design. The mean square quantization error is $\theta_n = \theta(\mathbb{Z}_n) := \mathbb{E}_X \min_{i=1, \dots, n} \|X - Z_i\|^2$. In the case where Z_1, \dots, Z_n are i.i.d. uniform on $\mathcal{C}_d(\delta)$, we will derive a simple approximation for the expected value of $\theta(\mathbb{Z}_n)$ and clearly demonstrate the δ -effect. Moreover, we will notice a strong similarity between efficient quantization designs and efficient designs constructed in Section 7.3.

7.1.5 Structure of the chapter and main results

In Section 7.2 we derive accurate approximations for the volume of intersection of an arbitrary d -dimensional cube with an arbitrary d -dimensional ball. These formulas will be heavily used in Section 7.3, which is the main section of the chapter dealing with the problem of $(1 - \gamma)$ -coverage of a cube by n balls. These approximations utilise the central limit theorem (CLT) and, as shall be demonstrated later, have the very appealing property of only depending on $\|Z\|$. One could investigate the use of exponential inequality type arguments, see Chapter 2 in [135], to develop approximations and bounds to the quantities of interest. The Hoeffding inequality used in Section 7.6.4 is an example of such exponential inequality. However, it does not seem that these inequalities will share the same appealing property the CLT approximations have, by depending only on $\|Z\|$. In Section 7.4 we extend some considerations of Section 7.3 to the problem of $(1 - \gamma)$ -coverage of the cube \mathcal{C}_d by smaller cubes. In Section 7.5 we argue that there is a strong similarity between efficient quantization designs and efficient designs of Section 7.3. In Appendix A, Section 7.6, we briefly mention several facts, used in the main part of the chapter, related to high-dimensional cubes and balls. In Appendix B, Section 7.7, we prove two simple but very important lemmas about distribution and moments of certain random variables.

The main contributions of this chapter are:

- an accurate approximation (7.2.16) for the volume of intersection of an arbitrary d -dimensional cube with an arbitrary d -dimensional ball;
- an accurate approximation (7.3.8) for the expected volume of intersection of the cube \mathcal{C}_d with n balls with uniform random centers $Z_j \in \mathcal{C}_d(\delta)$;
- closed-form expression of Section 7.4.2 for the expected volume of intersection the cube \mathcal{C}_d with n cubes with uniform random centers $Z_j \in \mathcal{C}_d(\delta)$;
- construction of efficient schemes of quantization and $(1 - \gamma)$ -coverage of the cube \mathcal{C}_d by n balls;
- large-scale numerical study.

7.2 Volume of intersection of a cube and a ball

7.2.1 The main quantity of interest

Consider the following problem. Let us take the cube $\mathcal{C}_d = [-1, 1]^d$ of volume $\text{vol}(\mathcal{C}_d) = 2^d$ and a ball $\mathcal{B}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| \leq r\}$ centered at a point $Z = (z_1, \dots, z_d) \in \mathbb{R}^d$; this point Z could be outside \mathcal{C}_d . Denote the fraction of the cube \mathcal{C}_d covered by the ball $\mathcal{B}_d(Z, r)$ by

$$C_{d,Z,r} = \text{vol}(\mathcal{C}_d \cap \mathcal{B}_d(Z, r)) / 2^d. \quad (7.2.1)$$

Our aim is to approximate $C_{d,Z,r}$ for arbitrary d , Z and r . To do this, we shall use CLT (Central Limit Theorem). We will derive a CLT-based normal approximation in

Section 7.2.3 and then, using an asymptotic expansion in the CLT for non-identically distributed r.v., we will improve this normal approximation in Section 7.2.4. In Section 7.6.8 we consider a more direct approach for approximating $C_{d,Z,r}$ based on the use of characteristic functions and the fact that $C_{d,Z,r}$ is a c.d.f. of $\|U - Z\|$, where $U = (u_1, \dots, u_d)$ is random vector with uniform distribution on \mathcal{C}_d . From this, $C_{d,Z,r}$ can be expressed through the convolution of one-dimensional c.d.f.'s. Using this approach we can evaluate the quantity $C_{d,Z,r}$ with high accuracy but the calculations are rather time-consuming. Moreover, entirely new computations have to be made for different Z and, therefore, the approximation of Section 7.2.4 is more appealing.

Note that in the special case $Z = 0$, several approximations for the quantity $C_{d,0,r}$ have been derived in [132] but their methods cannot be generalized to arbitrary Z .

7.2.2 A generalization of the quantity (7.2.1)

In the next sections, we will need another quantity which slightly generalizes (7.2.1). Assume that we have the cube $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$ of volume $\text{vol}(\mathcal{C}_d(\delta)) = (2\delta)^d$, the ball $\mathcal{B}_d(Z', r') = \{Y \in \mathbb{R}^d : \|Y - Z'\| \leq r'\}$ with a center at a point $Z' = (z'_1, \dots, z'_d)$. Denote the fraction of the cube $\mathcal{C}_d(\delta)$ covered by the ball $\mathcal{B}_d(Z', r')$ by

$$C_{d,Z',r'}^{(\delta)} = \text{vol}(\mathcal{C}_d(\delta) \cap \mathcal{B}_d(Z', r')) / (2\delta)^d. \quad (7.2.2)$$

Then, by changing the coordinates and the radius

$$Z = Z'/\delta = (z'_1/\delta, \dots, z'_d/\delta) \quad \text{and} \quad r = r'/\delta, \quad (7.2.3)$$

we obtain

$$C_{d,Z',r'}^{(\delta)} = C_{d,Z,r}. \quad (7.2.4)$$

7.2.3 Normal approximation for the quantity (7.2.1)

Let $U = (u_1, \dots, u_d)$ be a random vector with uniform distribution on \mathcal{C}_d so that u_1, \dots, u_d are i.i.d.r.v. uniformly distributed on $[-1, 1]$. Then for given $Z = (z_1, \dots, z_d) \in \mathbb{R}^d$ and any $r > 0$,

$$C_{d,Z,r} = \mathbb{P} \{ \|U - Z\| \leq r \} = \mathbb{P} \{ \|U - Z\|^2 \leq r^2 \} = \mathbb{P} \left\{ \sum_{j=1}^d (u_j - z_j)^2 \leq r^2 \right\}.$$

That is, $C_{d,Z,r}$, as a function of r , is the c.d.f. of the r.v. $\|U - Z\|$.

Let u have a uniform distribution on $[-1, 1]$ and $|z| \leq 1$. In view of Lemma 1 of Section 7.7, the density of the r.v. $\eta_z = (u - z)^2$ is

$$\varphi_z(t) = \begin{cases} 1/(2\sqrt{t}) & \text{for } 0 < t \leq (1 - |z|)^2 \\ 1/(4\sqrt{t}) & \text{for } (1 - |z|)^2 < t \leq (1 + |z|)^2 \\ 0 & \text{otherwise} \end{cases} \quad (7.2.5)$$

and

$$\mathbb{E}\eta_z = z^2 + \frac{1}{3}, \quad \text{var}(\eta_z) = \frac{4}{3} \left(z^2 + \frac{1}{15} \right), \quad \mu_z^{(3)} = \frac{16}{15} \left(z^2 + \frac{1}{63} \right), \quad (7.2.6)$$

where $\mu_z^{(3)}$ is the third central moment: $\mu_z^{(3)} = E[\eta_z - E\eta_z]^3$.

For $|z| > 1$, the density of $\eta_z = (u - z)^2$ is

$$\varphi_z(t) = \begin{cases} 1/(4\sqrt{t}) & \text{for } (1 - |z|)^2 < t \leq (1 + |z|)^2 \\ 0 & \text{otherwise} \end{cases} \quad (7.2.7)$$

with expressions (7.2.6) for $\mathbb{E}\eta_z$, $\text{var}(\eta_z)$ and $\mu_z^{(3)}$ not changing.

Consider the r.v.

$$\|U - Z\|^2 = \sum_{i=1}^d \eta_{z_i} = \sum_{j=1}^d (u_j - z_j)^2. \quad (7.2.8)$$

From (7.2.6), its mean is

$$\mu_{d,Z} = \mathbb{E}\|U - Z\|^2 = \|Z\|^2 + \frac{d}{3}. \quad (7.2.9)$$

Using independence of u_1, \dots, u_d , we also obtain from (7.2.6):

$$\sigma_{d,Z}^2 = \text{var}(\|U - Z\|^2) = \frac{4}{3} \left(\|Z\|^2 + \frac{d}{15} \right) \quad (7.2.10)$$

and

$$\mu_{d,Z}^{(3)} = \mathbb{E}[\|U - Z\|^2 - \mu_{d,Z}]^3 = \sum_{j=1}^d \mu_{z_j}^{(3)} = \frac{16}{15} \left(\|Z\|^2 + \frac{d}{63} \right). \quad (7.2.11)$$

If d is large enough then the conditions of the CLT for $\|U - Z\|^2$ are approximately met and the distribution of $\|U - Z\|^2$ is approximately normal with mean $\mu_{d,Z}$ and variance $\sigma_{d,Z}^2$. That is, we can approximate $C_{d,Z,r}$ by

$$C_{d,Z,r} \cong \Phi \left(\frac{r^2 - \mu_{d,Z}}{\sigma_{d,Z}} \right), \quad (7.2.12)$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution:

$$\Phi(t) = \int_{-\infty}^t \varphi(v) dv \quad \text{with } \varphi(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2}.$$

The approximation (7.2.12) has acceptable accuracy if $C_{d,Z,r}$ is not very small; for example, it falls inside a 2σ -confidence interval generated by the standard normal distribution; see Figures 7.1–7.2 as examples.

For a given β , let p_β satisfy $\Phi(\beta) = 1 - p_\beta$; for example, $p_\beta \simeq 0.025$ for $\beta = 2$. (β is the $1 - p_\beta$ quantile). As follows from (7.2.9), (7.2.10) and the approximation (7.2.12), we expect the approximate inequality $C_{d,Z,r} \gtrsim p_\beta$ if

$$r \geq R_{d,\|Z\|,\beta} = \left[\|Z\|^2 + d/3 - 2\beta \sqrt{\|Z\|^2/3 + d/45} \right]^{1/2}. \quad (7.2.13)$$

In many cases discussed in Section 7.3, the radius r does not satisfy the inequality (7.2.13) with $\beta = 2$ and even $\beta = 3$ and hence the normal approximation (7.2.12) is not satisfactorily accurate; this can be evidenced from Figures 7.1 – 7.16 below.

In the next section, we improve the approximation (7.2.12) by using an Edgeworth-type expansion in the CLT for sums of independent non-identically distributed r.v.

7.2.4 Improved normal approximation

General expansion in the central limit theorem for sums of independent non-identical r.v. has been derived by V. Petrov, see Theorem 7 in Chapter 6 in [88], see also Proposition 1.5.7 in [103]. The first three terms of this expansion have been specialized by V. Petrov in Section 5.6 in [89]. By using only the first term in this expansion, we obtain the following approximation for the distribution function of $\|U - Z\|^2$:

$$P\left(\frac{\|U - Z\|^2 - \mu_{d,Z}}{\sigma_{d,Z}} \leq x\right) \cong \Phi(x) + \frac{\mu_{d,Z}^{(3)}}{6(\sigma_{d,Z}^2)^{3/2}}(1 - x^2)\varphi(x),$$

leading to the following improved form of (7.2.12):

$$C_{d,Z,r} \cong \Phi(t) + \frac{\|Z\|^2 + d/63}{5\sqrt{3}(\|Z\|^2 + d/15)^{3/2}}(1 - t^2)\varphi(t), \quad (7.2.14)$$

where

$$t = t_{d,\|Z\|,r} = \frac{r^2 - \mu_{d,Z}}{\sigma_{d,Z}} = \frac{\sqrt{3}(r^2 - \|Z\|^2 - d/3)}{2\sqrt{\|Z\|^2 + d/15}}. \quad (7.2.15)$$

From the viewpoint of Section 7.3, the range of most important values of t from (7.2.15) is -3 ± 1 . For such values of t , the uncorrected normal approximation (7.2.12) significantly overestimates the values of $C_{d,Z,r}$, see Figures 7.1 – 7.16 below. The approximation (7.2.14) brings the normal approximation down and makes it much more accurate. The other terms in Petrov's expansion of [88] and [89] continue to bring the approximation down (in a much slower fashion) so that the approximation (7.2.14) still slightly overestimates the true value of $C_{d,Z,r}$ (at least, in the range of interesting values of t from (7.2.15)). However, if d is large enough (say, $d \geq 20$) then the approximation (7.2.14) is very accurate and no further correction is needed.

A very attractive feature of the approximations (7.2.12) and (7.2.15) is their dependence on Z through $\|Z\|$ only. We could have specialized for our case the next terms in Petrov's approximation but these terms no longer depend on $\|Z\|$ only (this fact can be verified from the formula (7.7.5) for the fourth moment of the r.v. $\nu_z = (z - u)^2$) and hence the next terms are much more complicated. Moreover, adding one or two extra terms from Petrov's expansion to the approximation (7.2.14) does not fix the problem entirely for all Z and r . Instead, the author proposes a slight adjustment to the r.h.s of (7.2.14) to improve this approximation, especially for small dimensions. Specifically, the author suggests the approximation

$$C_{d,Z,r} \cong \Phi(t) + c_d \frac{\|Z\|^2 + d/63}{5\sqrt{3}(\|Z\|^2 + d/15)^{3/2}}(1 - t^2)\varphi(t), \quad (7.2.16)$$

where $c_d = 1 + 3/d$ if the point Z lies on the diagonal of the cube \mathcal{C}_d and $c_d = 1 + 4/d$ for a typical (random) point Z . This small adjustment was chosen manually (and non theoretically) on the basis of making the approximation closer to simulation results; it is clear the adjustment should tend to zero as d increases, but the rate and the multiplying constant were chosen after trial and error.

For typical (random) points $Z \in \mathcal{C}_d$, the values of $C_{d,Z,r}$ are marginally smaller than for the points on the diagonal of \mathcal{C}_d having the same norm, but the difference is

very small. In addition to the points on the diagonal, there are other special points: the points whose components are all zero except for one. For such points, the values of $C_{d,Z,r}$ are smaller than for typical points Z with the same norm, especially for small r . Such points, however, are of no value for us as they are not typical and the author has never observed in simulations random points that come close to these truly exceptional points.

7.2.5 Simulation study

In Figures 7.1 – 7.16 we demonstrate the accuracy of approximations (7.2.12), (7.2.14) and (7.2.16) for $C_{d,Z,r}$ in dimensions $d = 10, 50$ for the following locations of Z :

- (i) $Z = 0$, the center of the cube \mathcal{C}_d ;
- (ii) $\|Z\| = \sqrt{d}$, with Z being a vertex of the cube \mathcal{C}_d ;
- (iii) Z lies on a diagonal of \mathcal{C}_d with $|z_j| = \lambda \geq 0$ for all $j = 1, \dots, d$ and $\|Z\| = \lambda\sqrt{d}$;
- (iv) Z is a random vector uniformly distributed on the sphere $\mathcal{S}_d(0, v)$ with some $v > 0$.

There are figures of two types. In the figures of the first type, we plot $C_{d,Z,r}$ over a wide range of r ensuring that values of $C_{d,Z,r}$ lie in the whole range $[0, 1]$. In the figures of the second type, we plot $C_{d,Z,r}$ over a much smaller range of r with $C_{d,Z,r}$ lying in the range $[0, \varepsilon]$ for some small positive ε such as $\varepsilon = 0.015$. For the purpose of using the approximations of Section 7.3, we need to assess the accuracy of all approximations for smaller values of $C_{d,Z,r}$ and hence the second type of plots are often more insightful. In Figures 7.1 – 7.14, the solid black line depicts values of $C_{d,Z,r}$ computed via Monte Carlo methods with 200,000 repetitions, the blue dashed, the red dot-dashed and green long dashed lines display approximations (7.2.12), (7.2.14) and (7.2.16), respectively.

In the case where Z is a random vector uniformly distributed on a sphere $\mathcal{S}_d(0, v)$, the style of the figures of the second type is slightly changed to adapt for this choice of Z and provide more information for Z which do or do not belong to the cube \mathcal{C}_d . In Figure 7.15 and Figure 7.16, the thick dashed red lines correspond to random points $Z \in \mathcal{S}_d(0, v) \cap \mathcal{C}_d$. The thick dot-dashed orange lines correspond to random points $Z \in \mathcal{S}_d(0, v)$ such that $Z \notin \mathcal{C}_d$. Approximations (7.2.12) and (7.2.14) are depicted in the same manner as previous figures but the approximation (7.2.16) is now represented by a solid green line. The thick solid red line displays values of $C_{d,Z,r}$ for Z on the diagonal of \mathcal{C}_d with $\|Z\| = v$ with $v = 1.5$ for $d = 10$ and $v = 1.75$ for $d = 50$.

From the simulations that led to Figures 7.1 – 7.16 we can make the following conclusions.

- The normal approximation (7.2.12) is quite satisfactory unless the value $C_{d,Z,r}$ is small.
- The accuracy of all approximations improves as d grows.

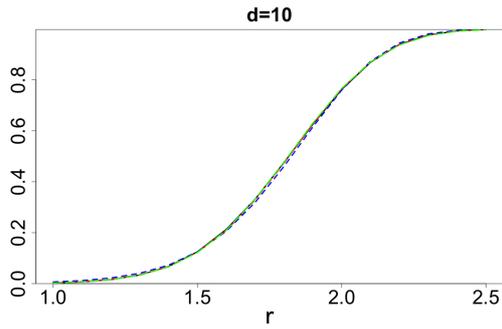


Figure 7.1: $d = 10$, $Z = 0$, $r \in [1, 2.5]$.

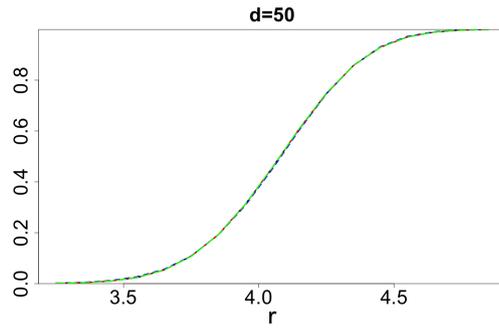


Figure 7.2: $d = 50$, $Z = 0$, $r \in [3.2, 4.9]$.

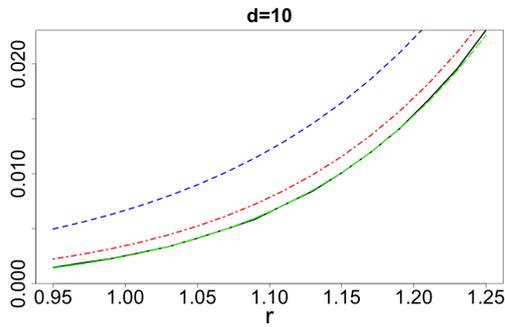


Figure 7.3: $d = 10$, $Z = 0$, $r \in [0.95, 1.25]$.

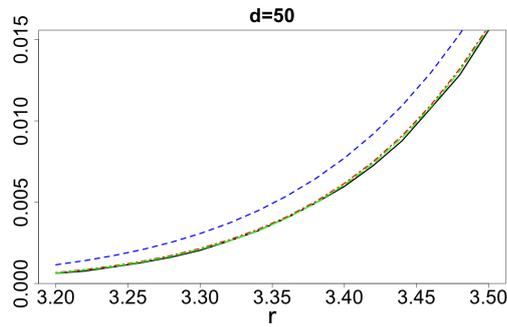


Figure 7.4: $d = 50$, $Z = 0$, $r \in [3.2, 3.5]$.

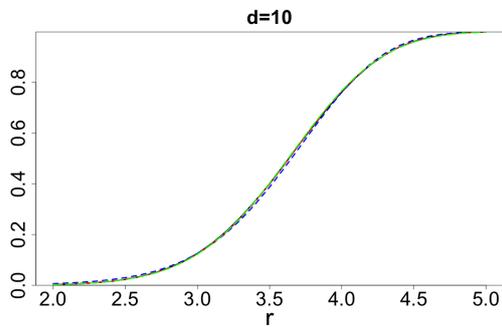


Figure 7.5: $d = 10$, Z is a vertex of \mathcal{C}_d , $r \in [2, 5]$.

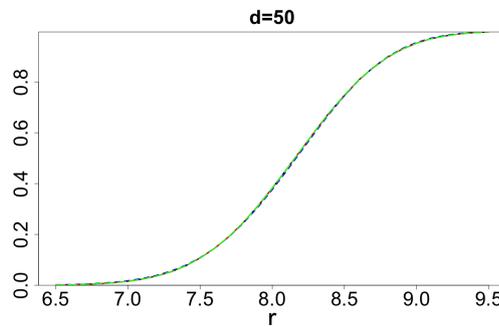


Figure 7.6: $d = 50$, Z is a vertex of \mathcal{C}_d , $r \in [6.5, 9.5]$.

- The approximation (7.2.16) is very accurate even if the values $C_{d,Z,r}$ are very small.
- If d is large enough then the approximations (7.2.14) and (7.2.16) are practically identical and are extremely accurate.

We remark that large values of $C_{d,Z,r}$ are of little interest here for the following reason. If the coverage from a single ball results in a large value of $C_{d,Z,r}$, then covering a cube by n of these balls (as we shall discuss in the next section) will result in an inefficient and trivial complete covering of the cube, where the value of r will be significantly larger than the smallest radius required to obtain full coverage. It is

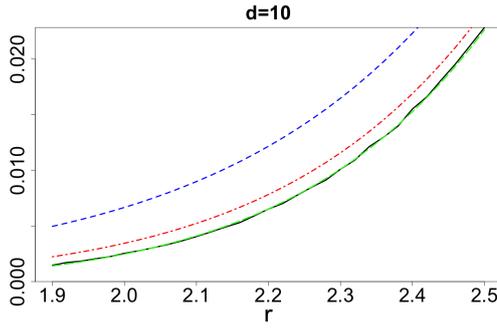


Figure 7.7: $d = 10$, Z is a vertex of \mathcal{C}_d , $r \in [1.9, 2.5]$.

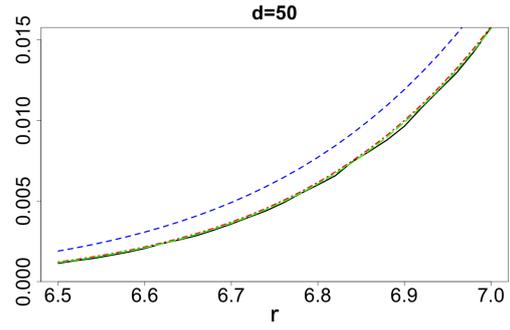


Figure 7.8: $d = 50$, Z is a vertex of \mathcal{C}_d , $r \in [6.5, 7]$.

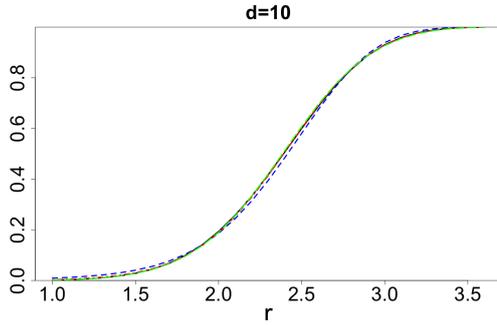


Figure 7.9: Z is at half-diagonal with $\|Z\| = \frac{1}{2}\sqrt{10}$

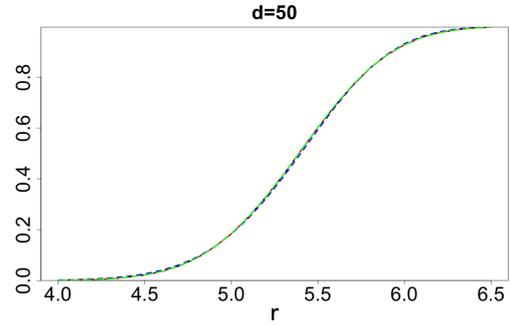


Figure 7.10: Z is at half-diagonal, $\|Z\| = \frac{1}{2}\sqrt{50}$

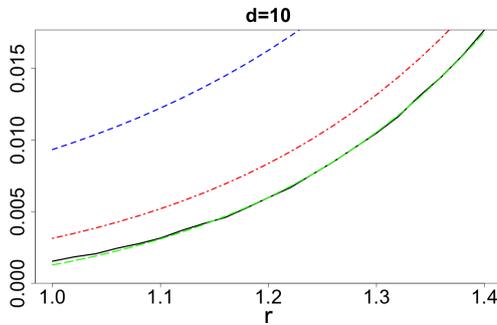


Figure 7.11: Z is at half-diagonal, $\|Z\| = \frac{1}{2}\sqrt{10}$

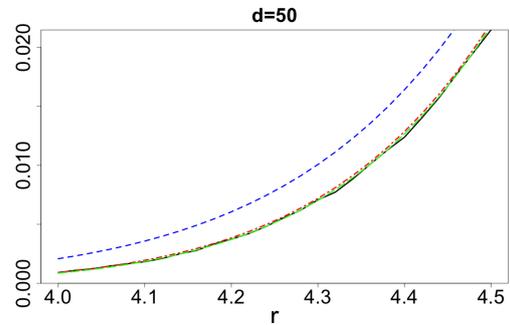


Figure 7.12: Z is at half-diagonal, $\|Z\| = \frac{1}{2}\sqrt{50}$

only when contributions from each individual ball is small e.g. small values of $C_{d,Z,r}$, that the problem of computing covering with n balls becomes non-trivial.

7.3 Covering a cube by n balls

In this section, we consider the main problem of covering the cube $\mathcal{C}_d = [-1, 1]^d$ by the union of n balls $\mathcal{B}_d(Z_j, r)$ as formulated in Section 7.1.2. We will discuss different schemes of choosing the set of ball centers $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ for given d and n . The radius r will then be chosen to achieve the required probability of covering: $C_d(\mathbb{Z}_n, r) \geq 1 - \gamma$. Most of the schemes will involve one or several parameters which

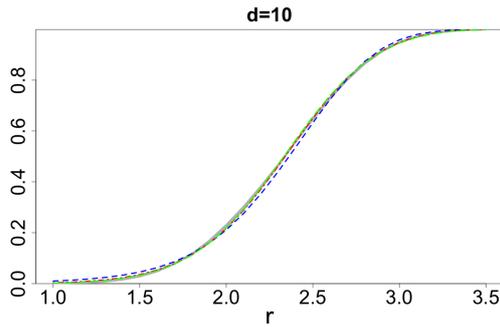


Figure 7.13: $d = 10$, $Z \in \mathcal{S}_{10}(0, 1.5)$, $r \in [1, 3.5]$

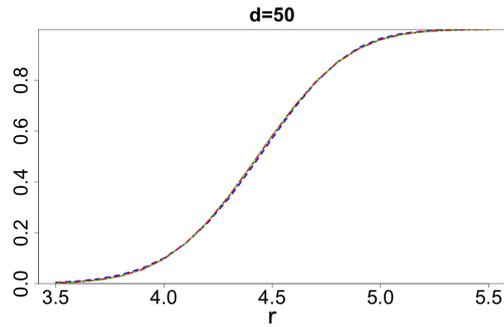


Figure 7.14: $d = 50$, $Z \in \mathcal{S}_{50}(0, 1.75)$, $r \in [3.5, 5.5]$

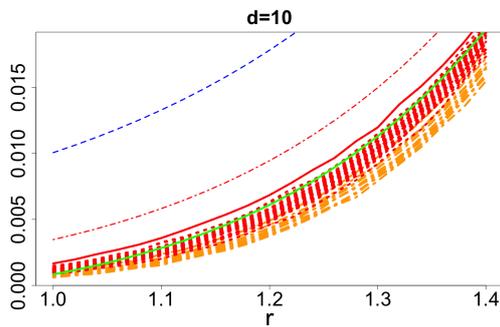


Figure 7.15: $d = 10$, $Z \in \mathcal{S}_{10}(0, 1.5)$, $r \in [1, 1.4]$

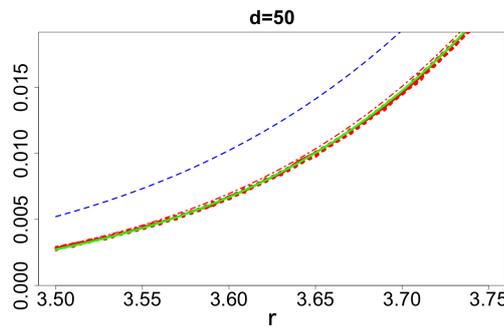


Figure 7.16: $d = 50$, $Z \in \mathcal{S}_{50}(0, 1.75)$, $r \in [3.5, 3.75]$

we will want to choose in an optimal way.

7.3.1 The main covering scheme

The following will be our main scheme for choosing $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$.

Scheme 1. Z_1, \dots, Z_n are *i.i.d.* random vectors uniformly distributed in the cube $C_d(\delta) = [-\delta, \delta]^d$, where $\delta \in [0, 1]$ is a parameter.

We will formulate several other covering schemes and compare them with Scheme 1. The reasons why Scheme 1 has been chosen as the main scheme are the following.

- It is easier to theoretically investigate than all other non-trivial schemes.
- It includes, as a special case when $\delta = 1$, the scheme which is very popular in practice of Monte-Carlo [79] and global random search [142, 146] and is believed to be rather efficient (this is not true).
- Numerical studies provided below show that Scheme 1 with optimal δ provides coverings which are rather efficient, especially for large d ; see Section 7.3.5 for a discussion regarding this issue.

7.3.2 Theoretical investigation of Scheme 1

Let Z_1, \dots, Z_n be i.i.d. random vectors uniformly distributed in the cube $\mathcal{C}_d(\delta)$ with $0 < \delta \leq 1$. Then, for given $U = (u_1, \dots, u_d) \in \mathbb{R}^d$ (we will later assume U is uniformly distributed in $[-1, 1]^d$):

$$\begin{aligned} \mathbb{P}\{U \in \mathcal{B}_d(\mathbb{Z}_n, r)\} &= 1 - \prod_{j=1}^n \mathbb{P}\{U \notin \mathcal{B}_d(Z_j, r)\} \\ &= 1 - \prod_{j=1}^n (1 - \mathbb{P}\{U \in \mathcal{B}_d(Z_j, r)\}) \\ &= 1 - \left(1 - \mathbb{P}_Z\{\|U - Z\| \leq r\}\right)^n, \end{aligned} \quad (7.3.1)$$

where $\mathcal{B}_d(\mathbb{Z}_n, r)$ is defined in (7.1.2). The main characteristic of interest $C_d(\mathbb{Z}_n, r)$, defined in (7.1.1), the proportion of the cube covered by the union of balls $\mathcal{B}_d(\mathbb{Z}_n, r)$, is simply

$$C_d(\mathbb{Z}_n, r) = \mathbb{E}_U \mathbb{P}\{U \in \mathcal{B}_d(\mathbb{Z}_n, r)\}, \quad (7.3.2)$$

where the expectation is taken with respect to the uniformly distributed $U \in [-1, 1]^d$. Continuing (7.3.1), note that

$$\mathbb{P}_Z\{\|U - Z\| \leq r\} = \mathbb{P}_Z\left\{\sum_{j=1}^d (z_j - u_j)^2 \leq r^2\right\} = C_{d,U,r}^{(\delta)}, \quad (7.3.3)$$

where $C_{d,U,r}^{(\delta)}$ is defined by the formula (7.2.2). From (7.2.3) and (7.2.4) we have $C_{d,U,r}^{(\delta)} = C_{d,U/\delta,r/\delta}$ where $C_{d,U/\delta,r/\delta}$ is the quantity defined by (7.2.1). This quantity can be approximated in a number of different ways as shown in Section 7.2. We will compare (7.2.12), the simplest of the approximations, with the approximation given in (7.2.16). Approximation (7.2.12) gives

$$C_{d,U,r}^{(\delta)} = C_{d,U/\delta,r/\delta} \cong \Phi\left(\frac{(r/\delta)^2 - \|U\|^2/\delta^2 - d/3}{2\sqrt{\|U\|^2/(3\delta^2) + d/45}}\right), \quad (7.3.4)$$

whereas approximation (7.2.16) provides

$$C_{d,U,r}^{(\delta)} \cong \Phi(t_\delta) + c_d \frac{\|U\|^2/\delta^2 + d/63}{5\sqrt{3}(\|U\|^2/\delta^2 + d/15)^{3/2}} (1 - t_\delta^2)\varphi(t_\delta), \quad (7.3.5)$$

with $c_d = 1 + 4/d$ and

$$t_\delta = \frac{(r/\delta)^2 - \|U\|^2/\delta^2 - d/3}{2\sqrt{\|U\|^2/(3\delta^2) + d/45}}.$$

From (7.6.6), $\mathbb{E}\|U\|^2 = d/3$ and $\text{var}(\|U\|^2) = 4d/45$. Moreover, if d is large enough then $\|U\|^2 = \sum_{j=1}^d u_j^2$ is approximately normal.

We shall simplify the expression (7.3.1) by using the approximation

$$(1 - t)^n \simeq e^{-nt}, \quad (7.3.6)$$

which is a good approximation for small values of t and moderate values of nt ; this agrees with the ranges of d , n and r we are interested in.

We can combine the expressions (7.3.2) and (7.3.1) with approximations (7.3.4), (7.3.5) and (7.3.6) as well as with the normal approximation for the distribution of $\|U\|^2$, to arrive at two final approximations for $C_d(\mathbb{Z}_n, r)$ that differ in complexity. If the original normal approximation of (7.3.4) is used then we obtain

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_1(s, r) \varphi(s) ds, \quad (7.3.7)$$

with

$$\psi_1(s, r) = \exp \{-n\Phi(c_{s,r})\}, \quad c_{s,r} = \frac{3(r/\delta)^2 - s' - d}{2\sqrt{s' + d/5}}, \quad s' = (d + 2s\sqrt{d/5})/\delta^2.$$

If approximation (7.3.5) is used, we obtain:

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_2(s, r) \varphi(s) ds, \quad (7.3.8)$$

with

$$\psi_2(s, r) = \exp \left\{ -n \left(\Phi(c_{s,r}) + \left(1 + \frac{4}{d} \right) \frac{s' + d/21}{5[s' + d/5]^{3/2}} (1 - c_{s,r}^2) \varphi(c_{s,r}) \right) \right\}.$$

7.3.3 Simulation study for assessing accuracy of approximations (7.3.7) and (7.3.8)

In Figures 7.17–7.22, $C_d(\mathbb{Z}_n, r)$ is represented by a solid black line and has been obtained via Monte Carlo methods with 50,000 repetitions. Approximation (7.3.7) is indicated by a dashed blue line and approximation (7.3.8) is represented by long dashed green lines. All figures demonstrate that approximation (7.3.8) is extremely accurate across different dimensions and values of n . This approximation is much superior to approximation (7.3.7). In these figures, the value of r has been chosen such that $C_d(\mathbb{Z}_n, r) \simeq 0.9$. The values of r used are included in the Figure captions.

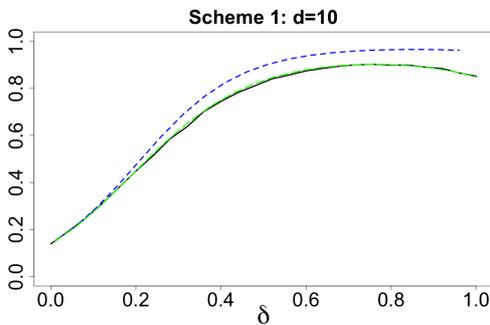


Figure 7.17: $C_d(\mathbb{Z}_n, r)$ and approximations: $n = 128, r = 1.520$.

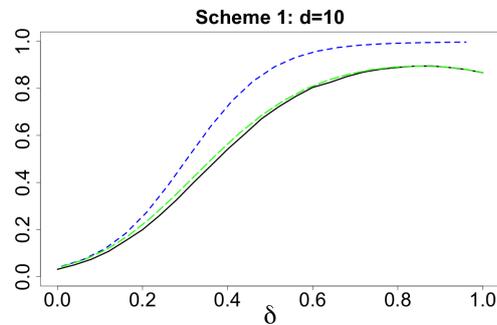


Figure 7.18: $C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 1.291$.

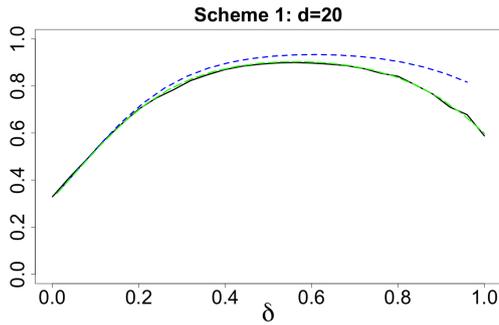


Figure 7.19: $C_d(\mathbb{Z}_n, r)$ and approximations: $n = 128, r = 2.460$.

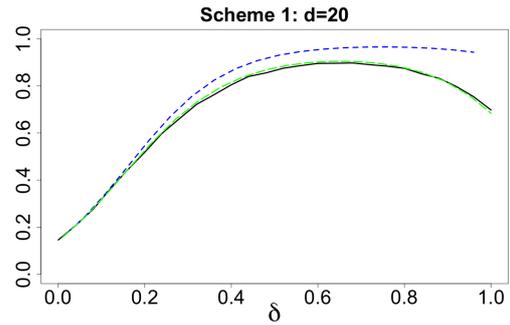


Figure 7.20: $C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 2.290$.

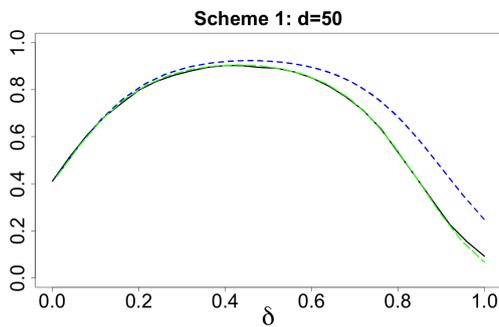


Figure 7.21: $C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 4.020$.

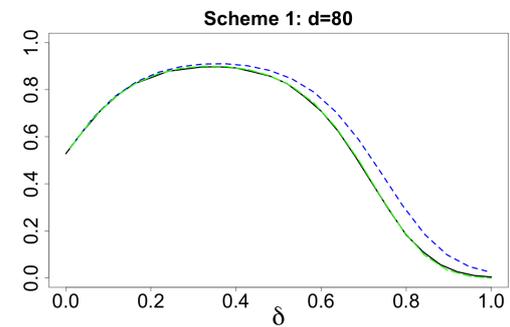


Figure 7.22: $C_d(\mathbb{Z}_n, r)$ and approximations: $n = 512, r = 5.175$.

7.3.4 Other schemes

In addition to Scheme 1, the following schemes for choosing $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ have also been considered.

Scheme 2. $Z_1 = 0$; Z_2, \dots, Z_n are *i.i.d.* random vectors uniformly distributed in the cube $C_d(\delta) = [-\delta, \delta]^d$.

Scheme 3. Z_1, \dots, Z_n are taken from the minimum-aberration fractional factorial design on vertices of the cube $C_d(\delta) = [-\delta, \delta]^d$. (See [14].)

Scheme 4. Z_1, \dots, Z_n are *i.i.d.* random vectors on $C_d(\delta)$ with independent components distributed according to Beta-distribution with density (7.6.3) with some $\alpha > 0$.

Scheme 5. Z_1, \dots, Z_n are *i.i.d.* random vectors uniformly distributed in the ball $\mathcal{B}_d(\delta)$.

Scheme 6. Z_1, \dots, Z_n are *i.i.d.* random vectors uniformly distributed on the sphere $\mathcal{S}_d(\delta)$.

Scheme 7. Z_1, \dots, Z_n are taken from a low-discrepancy Sobol's sequence on the cube $\mathcal{C}_d(\delta)$. (See [124, 125].)

The rationale behind the choice of these schemes is as follows. By studying Scheme 2, we test the importance of inclusion of 0 into \mathbb{Z}_n . The author propositioned that if we included 0 into \mathbb{Z}_n , the optimal value of δ may increase for some of the schemes making them more efficient; this effect has not been detected.

Scheme 3 with optimal δ is an obvious candidate for being the most efficient. Unlike all other schemes considered, Scheme 3 is only defined for the values of n of the form $n = 2^k$ with $k \leq d$.

By using Scheme 4, we test the possibility of improving Scheme 1 by changing the distribution of points in the cube $\mathcal{C}_d(\delta)$. The author has found that the effect of distribution is strong and smaller values of α lead to more efficient covering schemes. By choosing α small enough, like $\alpha = 0.1$, we achieve the average efficiency of the covering schemes which is rather close to the efficiency of Scheme 3. Tables 7.1–7.3 contain results obtained for Scheme 4 with $\alpha = 0.5$ and $\alpha = 1.5$; if $\alpha = 1$ then Scheme 4 becomes Scheme 1.

From Section 7.6.4, we know that for constructing efficient designs we have to somehow restrict the norms of Z_j 's. In Schemes 5 and 6, we are trying to do this in an alternative way to Schemes 1 and 4.

Scheme 7 is a natural improvement of Scheme 1. As a particular case with $\delta = 1$, it contains one of the best known low-discrepancy sequences and hence Scheme 7 with $\delta = 1$ serves as the main benchmark with which we compare other schemes. For construction, the author has used the R-implementation of the Sobol's sequences; it is based on [46].

For all the schemes excluding Scheme 3, the sequences $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ are nested so that $\mathbb{Z}_n \subset \mathbb{Z}_m$ for all $n < m$; using the terminology of [58], these schemes provide on-line coverings of the cube. Note that for the chosen values of n , Scheme 7 also has some advantage over other schemes considered. Indeed, despite Sobol's sequences are nested, the values n of the form $n = 2^k$ are special for the Sobol's sequences and for such values of n the Sobol's sequences possess extra uniformity properties that they do not possess for other values of n .

7.3.5 Numerical comparison of schemes

In Tables 7.1–7.3, for Schemes 1,2,4,5,6 we present the smallest values of r required to achieve an 0.9-coverage on average. For these schemes, the value inside the brackets shows the value of δ required to obtain 0.9-coverage. For Schemes 3 and 7, we give the smallest value of r needed for a 0.9-coverage. To determine the values of r and δ , we have used Monte Carlo simulations with 50,000 iterations (note for Scheme 1 we could use Approximation (7.3.8) which is very accurate). To compute Scheme 3, we have used the R-package "FrF2".

In Figures 7.23–7.30 we plot $C_d(\mathbb{Z}_n, r)$ as a functions of $\delta \in [0, 1]$ across a number schemes, n and d . For these plots we have used the values of r provided in Tables 7.1–7.3 such that for Figures 7.23–7.26 which correspond to Scheme 1 and Scheme 2, the maximum coverage is very close to 0.9 and the optimal δ is very close to the values

presented in Tables 7.1–7.3. For Figures 7.27–7.30 the maximum coverage 0.9 is attained with δ provided in Tables 7.1–7.3. In Figures 7.23–7.30 the solid green line, long dashed red line, dashed blue line and dot dashed orange line correspond to $n = 64, 128, 512, 1024$ respectively. The vertical lines on these plots indicate the value of δ where the maximum coverage is obtained.

$d = 10$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Scheme 1	1.632 (0.70)	1.520 (0.78)	1.291 (0.86)	1.195 (0.90)
Scheme 1, $\delta = 1$	1.720 (1.00)	1.577 (1.00)	1.319 (1.00)	1.215 (1.00)
Scheme 2	1.634 (0.70)	1.520 (0.78)	1.291 (0.86)	1.195 (0.90)
Scheme 3	1.530 (0.44)	1.395 (0.48)	1.115 (0.50)	1.075 (0.50)
Scheme 4, $\alpha = 0.5$	1.629 (0.58)	1.505 (0.65)	1.270 (0.72)	1.165 (0.75)
Scheme 4, $\alpha = 1.5$	1.635 (0.80)	1.525 (0.88)	1.310 (1.00)	1.210 (1.00)
Scheme 5	1.645 (1.40)	1.530 (1.50)	1.330 (1.75)	1.250 (1.75)
Scheme 6	1.642 (1.25)	1.532 (1.35)	1.330 (1.50)	1.250 (1.70)
Scheme 7	1.595 (0.72)	1.485 (0.80)	1.280 (0.85)	1.170 (0.88)
Scheme 7, $\delta = 1$	1.678 (1.00)	1.534 (1.00)	1.305 (1.00)	1.187 (1.00)

Table 7.1: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 10$.

$d = 20$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Scheme 1	2.545 (0.50)	2.460 (0.55)	2.290 (0.68)	2.205 (0.70)
Scheme 1, $\delta = 1$	2.840 (1.00)	2.702 (1.00)	2.444 (1.00)	2.330 (1.00)
Scheme 2	2.545 (0.50)	2.460 (0.55)	2.290 (0.68)	2.205 (0.70)
Scheme 3	2.490 (0.32)	2.410 (0.35)	2.220 (0.40)	2.125 (0.44)
Scheme 4, $\alpha = 0.5$	2.540 (0.44)	2.455 (0.48)	2.285 (0.55)	2.220 (0.60)
Scheme 4, $\alpha = 1.5$	2.545 (0.60)	2.460 (0.65)	2.290 (0.76)	2.215 (0.78)
Scheme 5	2.550 (1.40)	2.467 (1.60)	2.305 (1.75)	2.235 (1.90)
Scheme 6	2.550 (1.40)	2.467 (1.58)	2.305 (1.75)	2.235 (1.90)
Scheme 7	2.520 (0.50)	2.445 (0.60)	2.285 (0.68)	2.196 (0.72)
Scheme 7, $\delta = 1$	2.750 (1.00)	2.656 (1.00)	2.435 (1.00)	2.325 (1.00)

Table 7.2: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 20$.

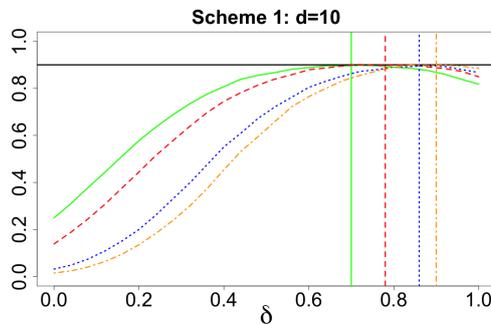


Figure 7.23: Scheme 1: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$

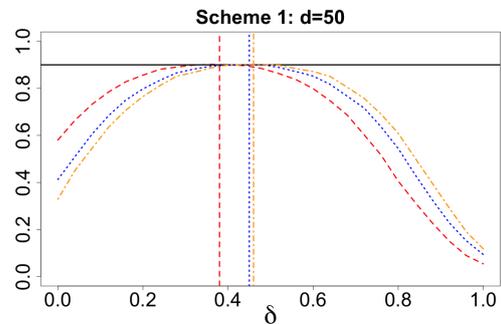


Figure 7.24: Scheme 1: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$

$d = 50$			
	$n = 128$	$n = 512$	$n = 1024$
Scheme 1	4.130 (0.38)	4.020 (0.45)	3.970 (0.46)
Scheme 1, $\delta = 1$	4.855 (1.00)	4.625 (1.00)	4.520 (1.00)
Scheme 2	4.130 (0.38)	4.020 (0.45)	3.970 (0.46)
Scheme 3	4.110 (0.21)	4.000 (0.25)	3.950 (0.28)
Scheme 4 $\alpha = 0.5$	4.130 (0.30)	4.020 (0.36)	3.970 (0.40)
Scheme 4 $\alpha = 1.5$	4.130 (0.42)	4.020 (0.48)	3.970 (0.52)
Scheme 5	4.130 (1.50)	4.020 (1.75)	3.970 (2.00)
Scheme 6	4.130 (1.50)	4.020 (1.75)	3.970 (2.00)
Scheme 7	4.115 (0.40)	4.015 (0.45)	3.965 (0.47)
Scheme 7, $\delta = 1$	4.395 (1.00)	4.379 (1.00)	4.366 (1.00)

Table 7.3: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 50$.

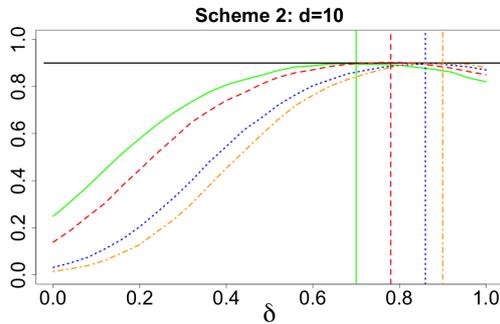


Figure 7.25: Scheme 2: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$

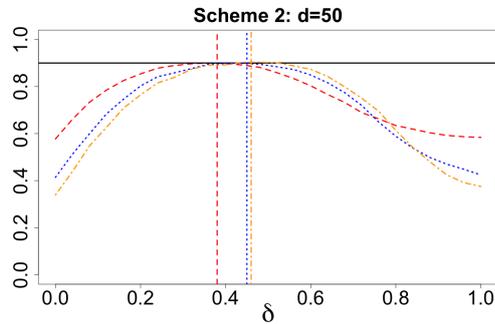


Figure 7.26: Scheme 2: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$

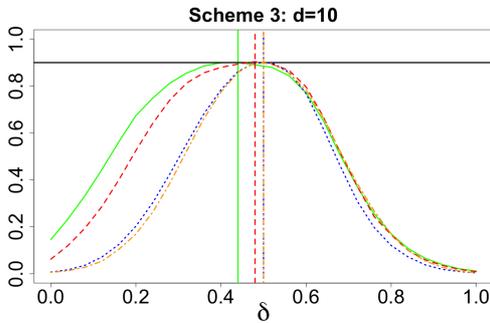


Figure 7.27: Scheme 3: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$

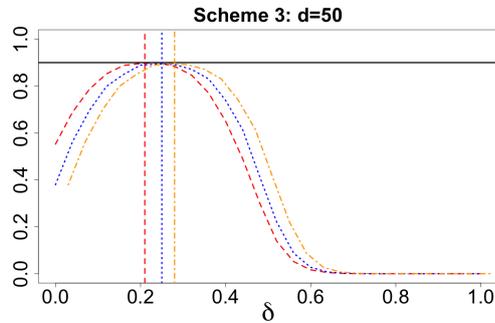


Figure 7.28: Scheme 3: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$

From Tables 7.1–7.3 and Figures 7.23–7.30 we arrive at the following conclusions:

- the δ -effect is very important and getting much stronger as d increases; the existence of the δ -effect is in agreement with the result of Section 7.6.4 describing the area of volume concentration in the cube \mathcal{C}_d .
- coverage of unadjusted low-discrepancy sequences is extremely low;
- properly δ -tuned deterministic Scheme 3, which uses fractional factorial designs of minimum aberration, provides excellent covering;

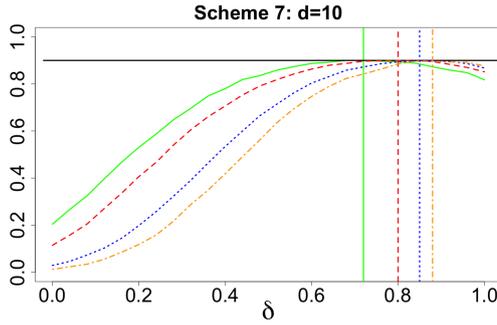


Figure 7.29: Scheme 7: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 10$

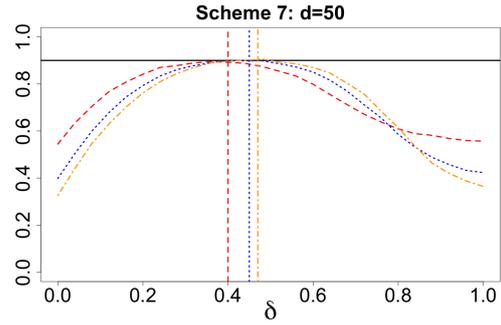


Figure 7.30: Scheme 7: $C_d(\mathbb{Z}_n, r)$ across δ for $d = 50$

- randomized Scheme 4 with suitably chosen parameters of the Beta-distribution, also provides very high quality coverage (on average);

7.4 Covering a cube by cubes

7.4.1 Volume of intersection of two cubes

Let us take two cubes: $\mathcal{C}_d = [-1, 1]^d$ and $\mathcal{C}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\|_\infty \leq r\}$, a cube of side length $2r$ centered at a point $Z = (z_1, \dots, z_d) \in \mathcal{C}_d$. Denote the fraction of the cube \mathcal{C}_d covered by $\mathcal{C}_d(Z, r)$ by

$$F_{d,Z,r} = \text{vol}(\mathcal{C}_d \cap \mathcal{C}_d(Z, r)) / 2^d. \quad (7.4.1)$$

Let, like in Section 7.2.3, $U = (u_1, \dots, u_d)$ be a random vector with uniform distribution on \mathcal{C}_d so that u_1, \dots, u_d are i.i.d.r.v. uniformly distributed on $[-1, 1]$. Then

$$F_{d,Z,r} = \mathbb{P}\{\|U - Z\|_\infty \leq r\} = \mathbb{P}\left\{\max_{1 \leq j \leq d} |u_j - z_j| \leq r\right\}.$$

That is, $F_{d,Z,r}$, as a function of r , is the c.d.f. of the r.v. $\|U - Z\|_\infty = \max_{1 \leq j \leq d} |u_j - z_j|$.

From Lemma 2 of Section 7.7 the c.d.f. of the r.v. $|u_j - z_j|$ is

$$G_{d,z_j}(t) = \mathbb{P}\{|u_j - z_j| \leq t\} = \begin{cases} 0 & \text{for } t \leq 0 \\ t & \text{for } 0 < t < 1 - |z_j| \\ \frac{1}{2}(1+t-|z_j|) & \text{for } 1 - |z_j| \leq t \leq 1 + |z_j| \\ 1 & \text{for } 1 + |z_j| < t. \end{cases}$$

Since the c.d.f. of a maximum of independent r.v. is the product of marginal c.d.f.'s, we obtain

$$F_{d,Z,r} = \prod_{j=1}^d G_{d,z_j}(r).$$

Two extreme particular cases of location of Z are:

- (i) $Z = 0$: $F_{d,0,r} = r^d$, $0 \leq r \leq 1$;

(ii) $\|Z\| = \sqrt{d}$, when Z is a vertex of the cube \mathcal{C}_d : $F_{d,V,r} = (r/2)^d$, $0 \leq r \leq 2$.

Assume now that we have the cube $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$ of volume $(2\delta)^d$ and another cube $\mathcal{C}_d(Z', r') = \{Y \in \mathbb{R}^d : \|Y - Z'\|_\infty \leq r'\}$ with a center at a point $Z' = (z'_1, \dots, z'_d)$. Denote the fraction of the cube $\mathcal{C}_d(\delta)$ covered by $\mathcal{C}_d(Z', r')$ by

$$F_{d,Z',r'}^{(\delta)} = \text{vol}(\mathcal{C}_d(\delta) \cap \mathcal{C}_d(Z', r')) / (2\delta)^d.$$

Then by changing the coordinates and the radius using (7.2.3) we get $F_{d,Z',r'}^{(\delta)} = F_{d,Z'/\delta,r'/\delta}$.

7.4.2 Proportion of a cube covered by smaller cubes with random centers

Let us take the cube $\mathcal{C}_d = [-1, 1]^d$ and n smaller cubes $\mathcal{C}_d(Z_j, r) = \{Y \in \mathbb{R}^d : \|Y - Z_j\|_\infty \leq r\}$ with centers at points $Z_j \in \mathbb{R}^d$. Denote the fraction of the cube \mathcal{C}_d covered by $\mathcal{C}_d(\mathbb{Z}_n, r) = \cup_{j=1}^n \mathcal{C}_d(Z_j, r)$, the union of these cubes, by

$$C_{d,\mathbb{Z}_n,r} = \text{vol}(\mathcal{C}_d \cap \mathcal{C}_d(\mathbb{Z}_n, r)) / 2^d.$$

Our aim is to obtain a closed form expression for this quantity for arbitrary d, r and n in the case when Z_1, \dots, Z_n are i.i.d. random vectors uniformly distributed in the cube $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$ with $0 < \delta \leq 1$.

Similarly to the combination of (7.3.1) with (7.3.3), for a given $U = (u_1, \dots, u_d) \in \mathbb{R}^d$,

$$\mathbb{P}\{U \in \mathcal{C}_d(\mathbb{Z}_n, r)\} = 1 - \left(1 - F_{d,U/\delta,r/\delta}\right)^n.$$

Similarly to (7.3.2),

$$C_{d,\mathbb{Z}_n,r} = \mathbb{E}_U \mathbb{P}\{U \in \mathcal{C}_d(\mathbb{Z}_n, r)\} = 1 - \mathbb{E}_U \left(1 - F_{d,U/\delta,r/\delta}\right)^n.$$

For an integer k , set

$$I_k = \frac{1}{2} \int_{-1}^1 [G_{d,u/\delta}(r/\delta)]^k du. \quad (7.4.2)$$

Then, using the binomial theorem, we have

$$C_{d,\mathbb{Z}_n,r} = 1 - \sum_{k=0}^n (-1)^k \binom{n}{k} I_k^d. \quad (7.4.3)$$

It is possible to evaluate (7.4.2) explicitly. For $k = 0$ and for $r \geq \delta + 1$, we clearly have $I_k = 1$. For $k \geq 1$ and $0 \leq r \leq \delta + 1$, the integral I_k takes different forms depending on the values of r and δ :

$$I_k = \begin{cases} (\delta - r) \left(\frac{r}{\delta}\right)^k - \frac{2\delta}{(k+1)} \left\{ \left(\frac{\delta+r-1}{2\delta}\right)^{k+1} - \left(\frac{r}{\delta}\right)^{k+1} \right\} & \text{for } r \leq \delta \\ (r - \delta) - \frac{2\delta}{(k+1)} \left\{ \left(\frac{\delta+r-1}{2\delta}\right)^{k+1} - 1 \right\} & \text{for } 0 \leq r - \delta \leq 1, r + \delta \geq 1 \\ (r - \delta) + 2\delta/(k+1) & \text{for } 0 \leq r - \delta \leq 1, r + \delta \leq 1. \end{cases}$$

In Figures 7.31–7.32, we depict values of $C_{d, \mathbb{Z}_n, r}$ (computed using (7.4.3)) as a function of δ for a number of choices of r . As in Section 7.3.5, we note that the δ -effect holds for the problem of coverage of the cube by smaller cubes.

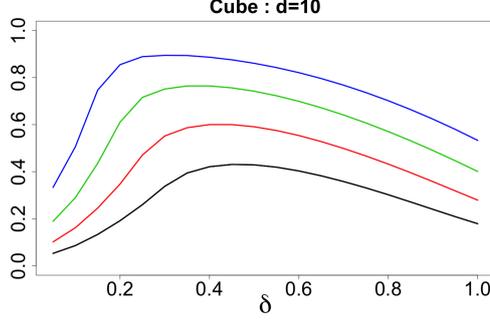


Figure 7.31: $n = 50$, $r \in [0.7, 0.85]$ increasing by 0.05.

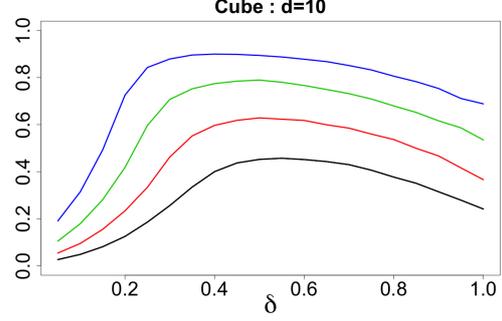


Figure 7.32: $n = 128$, $r \in [0.6, 0.8]$ increasing by 0.05.

7.5 Quantization

In this section, we briefly consider the following problem of quantization also known as the problem of facility location. Let $X = (x_1, \dots, x_d)$ be uniform on $\mathcal{C}_d = [-1, 1]^d$ and $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ be an n -point design. The mean square quantization error is $\theta_n = \theta(\mathbb{Z}_n) = \mathbb{E}_X \min_{i=1, \dots, n} \|X - Z_i\|^2$. In the case where Z_1, \dots, Z_n are i.i.d. uniform on $\mathcal{C}_d(\delta)$, we will derive a simple approximation for the expected value of $\theta(\mathbb{Z}_n)$ in order to demonstrate the δ -effect. We shall also notice a strong correlation in design efficiency used for quantization and for $(1 - \gamma)$ -covering as studied in Section 7.3.

The two characteristics, $C_d(\mathbb{Z}_n, r)$ and $\theta(\mathbb{Z}_n)$, are related as follows:

$$\begin{aligned} \mathbb{E}_{\mathbb{Z}_n} \theta(\mathbb{Z}_n) &= \mathbb{E}_{\mathbb{Z}_n} \mathbb{E}_X \min_{i=1, \dots, n} \|X - Z_i\|^2 = \mathbb{E}_X \mathbb{E}_{\mathbb{Z}_n} \min_{i=1, \dots, n} \|X - Z_i\|^2 \\ &= \int_{r \geq 0} r^2 d \mathbb{E}_{\mathbb{Z}_n} C_d(\mathbb{Z}_n, r). \end{aligned} \quad (7.5.1)$$

The relationship between quantization error and $C_d(\mathbb{Z}_n, r)$ is discussed in more detail in Chapter 8, see Section 8.1.3.

Using approximation (7.3.8) we obtain

$$\begin{aligned} \frac{d}{dr} (\mathbb{E}_{\mathbb{Z}_n} C_d(\mathbb{Z}_n, r)) &\cong f_\delta(r) := \frac{n \cdot r}{\delta} \int_{-\infty}^{\infty} \frac{\varphi(s) \varphi(c_{s,r}) \psi_2(s, r)}{\sqrt{s' + k}} \times \\ &\times \left[\sqrt{3} + \left(1 + \frac{4}{d}\right) \frac{\left(s' + \frac{d\delta^2}{63}\right)}{5(s' + k)^{3/2}} \left\{ \delta(c_{s,r}^3 - c_{s,r}) - \frac{\sqrt{3}(r^2 - \frac{d\delta^2}{3} - s')}{\sqrt{s' + k}} \right\} \right] ds \end{aligned}$$

with

$$\psi_2(s, r) = \exp \left\{ -n \left(\Phi(c_{s,r}) + \left(1 + \frac{4}{d}\right) \frac{\delta \left[s' + \frac{d\delta^2}{63} \right]}{5\sqrt{3} [s' + k]^{3/2}} (1 - c_{s,r}^2) \phi(c_{s,r}) \right) \right\}$$

and

$$c_{s,r} = \frac{\sqrt{3} \left(r^2 - s' - \frac{d\delta^2}{3} \right)}{2\delta\sqrt{s'+k}}, \quad s' = s\sqrt{\frac{4d}{45}} + d/3, \quad k = \frac{d\delta^2}{15}.$$

Therefore using relation (7.5.1) the approximation for $\mathbb{E}_{\mathbb{Z}_n}\theta(\mathbb{Z}_n)$ for Scheme 1 is:

$$\mathbb{E}\theta_n = \mathbb{E}_{\mathbb{Z}_n}\theta(\mathbb{Z}_n) \cong \int_{r \geq 0} r^2 f_\delta(r) dr. \quad (7.5.2)$$

In Figures 7.33–7.36, we assess the accuracy of the approximation (7.5.2). In these figures, the solid black line corresponds to $\mathbb{E}\theta_n$ obtained via Monte Carlo methods with 50,000 runs and the dashed red line depicts the approximation. We see that the accuracy of approximation (7.5.2) is very high for small n and large d . However, as Figure 7.34 shows, if d is not large enough but n is large, then the errors accumulate and the resulting approximation may not be accurate enough.

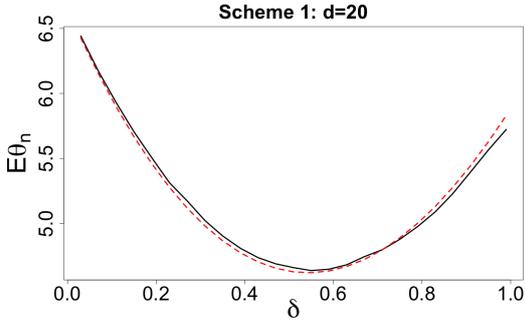


Figure 7.33: $\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 128$.

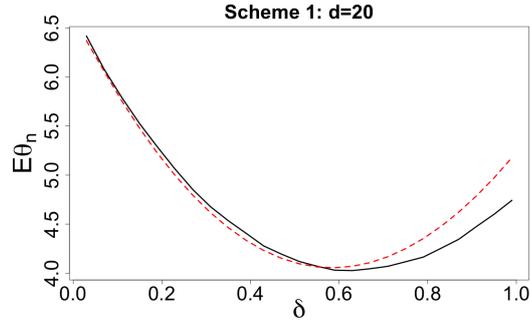


Figure 7.34: $\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 512$.

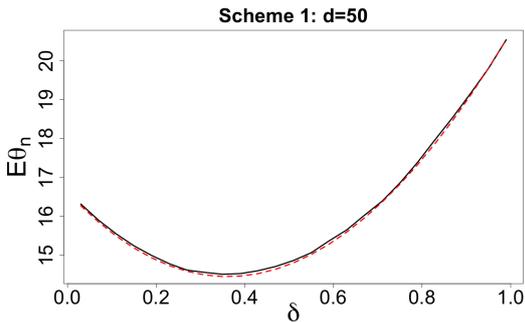


Figure 7.35: $\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 128$.

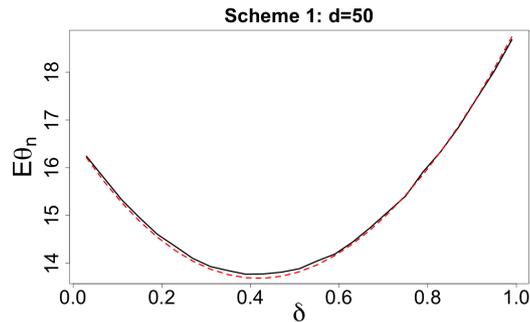


Figure 7.36: $\mathbb{E}\theta_n$ and approximation (7.5.2); $n = 512$.

As follows from results of [79, Ch.6], for efficient covering schemes the order of convergence of the covering radius to 0 as $n \rightarrow \infty$ is $n^{-1/d}$. Therefore, for the mean squared distance (which is the quantization error) we should expect the order $n^{-2/d}$ as $n \rightarrow \infty$. Therefore, for sake of comparison of quantization errors θ_n across n we renormalize this error from $\mathbb{E}\theta_n$ to $n^{2/d}\mathbb{E}\theta_n$.

In Tables 7.4–7.6, we present the minimum value of $n^{2/d}\mathbb{E}\theta_n$ for a selection of the schemes among those considered in Section 7.3. These values have been obtained via Monte Carlo methods with 50,000 iterations (note for Scheme 1 we could have used the accurate approximation given in (7.5.2)). In these tables, the value within the brackets corresponds to the value of δ where the minimum of $n^{2/d}\mathbb{E}\theta_n$ was obtained. For Scheme 3, typical behaviour of $\mathbb{E}\theta_n$ across δ for a number and n and d is presented in Figures 7.37–7.40.

$d = 10$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Scheme 1	4.153 (0.68)	4.105 (0.72)	3.992 (0.80)	3.925 (0.84)
Scheme 3	3.663 (0.40)	3.548 (0.44)	3.221 (0.48)	3.348 (0.52)
Scheme 4, $\alpha = 0.5$	4.072 (0.56)	4.013 (0.60)	3.839 (0.68)	3.770 (0.69)
Scheme 7	3.998 (0.68)	3.973 (0.76)	3.936 (0.80)	3.834 (0.82)
Scheme 7, $\delta = 1$	4.569 (1.00)	4.425 (1.00)	4.239 (1.00)	4.094 (1.00)

Table 7.4: Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across schemes and n for $d = 10$.

$d = 20$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Scheme 1	7.552 (0.52)	7.563 (0.56)	7.528 (0.64)	7.484 (0.68)
Scheme 3	7.298 (0.32)	7.270 (0.33)	7.133 (0.36)	7.016 (0.40)
Scheme 4, $\alpha = 0.5$	7.541 (0.40)	7.515 (0.44)	7.457 (0.52)	7.421 (0.54)
Scheme 7	7.445 (0.48)	7.464 (0.56)	7.487 (0.64)	7.453 (0.66)
Scheme 7, $\delta = 1$	9.089 (1.00)	9.133 (1.00)	8.87 (1.00)	8.681 (1.00)

Table 7.5: Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across schemes and n for $d = 20$.

$d = 50$			
	$n = 128$	$n = 512$	$n = 1024$
Scheme 1	17.608 (0.36)	17.634 (0.40)	17.643 (0.44)
Scheme 3	17.483 (0.20)	17.511 (0.24)	17.554 (0.27)
Scheme 4, $\alpha = 0.5$	17.590 (0.28)	17.670 (0.36)	17.620 (0.38)
Scheme 7, $\delta = 1$	20.196 (1.00)	21.231 (1.00)	21.711 (1.00)

Table 7.6: Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across schemes and n for $d = 50$.

We make the following two main conclusions from analyzing results of this numerical study:

- (a) the presence of a strong δ -effect, very similar to the effect observed in Section 7.3, and
- (b) for a given design \mathbb{Z}_n , there is a very strong correlation between the covering probability as studied in Section 7.3 and the normalized quantization error $n^{2/d}\mathbb{E}\theta(\mathbb{Z}_n)$.

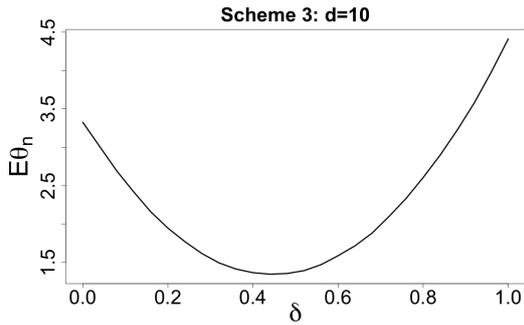


Figure 7.37: $\mathbb{E}\theta_n$ with $n = 128$.

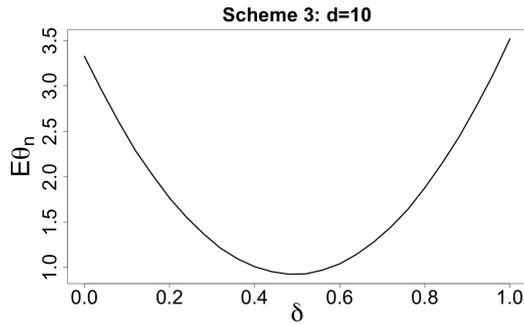


Figure 7.38: $\mathbb{E}\theta_n$ with $n = 512$.

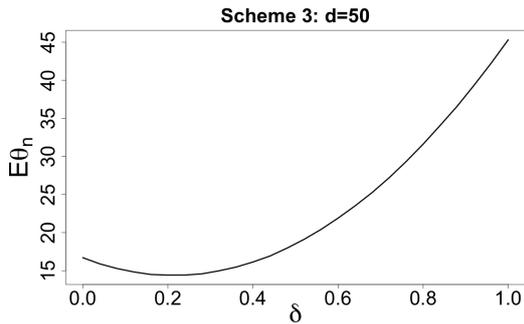


Figure 7.39: $\mathbb{E}\theta_n$ with $n = 128$.

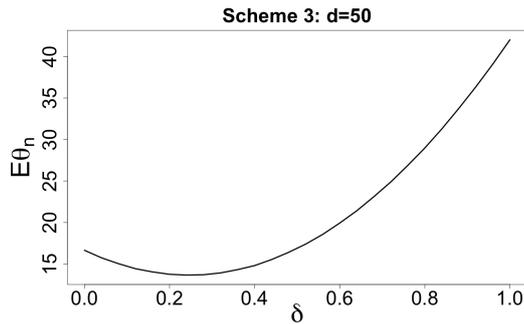


Figure 7.40: $\mathbb{E}\theta_n$ with $n = 512$.

By comparing the values of δ in Tables 7.4–7.6 with Tables 7.1–7.3, we see a strong similarity between efficient quantization schemes and efficient covering schemes.

7.6 Appendix A: Several facts about d -dimensional balls and cubes

In this appendix, we briefly consider several facts, used in the main part of the chapter, related to high-dimensional cubes and balls. Many of these facts are somewhat counter-intuitive and often lead to creation of wrong heuristics in multivariate optimization and misunderstanding of the behaviour of even simple algorithms in high-dimensional spaces. For more details concerning the material of Sections 7.6.1–7.6.4, see [12].

7.6.1 Volume of the ball

The volume of the ball $\mathcal{B}_d(r) = \{x \in \mathbb{R}^d : \|x\| \leq r\}$ can be computed by the formula

$$\text{vol}(\mathcal{B}_d(r)) = r^d V_d, \text{ where } V_d = \text{vol}(\mathcal{B}_d(1)) = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)}. \quad (7.6.1)$$

The volumes V_d decrease very fast as d grows. For example, $V_{100} \simeq 2.368 \cdot 10^{-40}$. As $d \rightarrow \infty$,

$$V_d^{1/d} \simeq \sqrt{2\pi e} \frac{1}{\sqrt{d}} + O\left(\frac{\log d}{d^{3/2}}\right). \quad (7.6.2)$$

7.6.2 Radius of the ball of unit volume

Define r_d by $\text{vol}(\mathcal{B}_d(r_d)) = 1$. Table 7.7 gives approximate values of r_d .

d	1	2	3	4	5	6	7	8	9
r_d	0.5	0.564	0.62	0.671	0.717	0.761	0.8	0.839	0.876
d	10	20	30	40	50	100	200	500	1000
r_d	0.911	1.201	1.43	1.626	1.8	2.49	3.477	5.45	7.682

Table 7.7: Radius of the ball of unit volume for different dimensions

From (7.6.2), for large d we have

$$r_d = \frac{\sqrt{d}}{\sqrt{2\pi e}} + O\left(\frac{1}{\sqrt{d}}\right),$$

where $1/\sqrt{2\pi e} \simeq 0.242$. This is only about twice smaller than $\sqrt{d}/2$, the length of the half-diagonal of the d -dimensional unit cube $[0, 1]^d$.

For $r_{d,2\delta}$ defined by $\text{vol}(\mathcal{B}_d(r_{d,2\delta})) = \text{vol}(\mathcal{C}_d(\delta)) = (2\delta)^d$, we have $r_{d,2\delta} = 2\delta r_d$.

7.6.3 Almost all the volume is near the boundary

First, consider the cube $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$, with $0 < \delta < 1$, as interior to the cube $\mathcal{C}_d = [-1, 1]^d$. For the ratio of the volumes of these two cubes, we have $\text{vol}(\mathcal{C}_d(\delta))/\text{vol}(\mathcal{C}_d) = \delta^d$ which tends to 0 (as $d \rightarrow \infty$) exponentially fast for any $\delta \in (0, 1)$.

If, as $d \rightarrow \infty$, δ changes getting closer to 1 but $1 - \delta$ tends to 0 slower than $1/d$, then the ratio of the two volumes still tends to 0. In particular, if $1 - \delta = c/d^{1-\delta}$ with $0 < \delta < 1$ then

$$\frac{\text{vol}(\mathcal{C}_d(\delta))}{\text{vol}(\mathcal{C}_d)} = \delta^d \simeq \exp\{-cd^{1-\delta}\} \rightarrow 0, \quad d \rightarrow \infty.$$

Consider now the balls $\mathcal{B}_d(1)$ and $\mathcal{B}_d(1 - \epsilon)$. The difference $\mathcal{B}_d(1) \setminus \mathcal{B}_d(1 - \epsilon)$ is called the annulus. Using (7.6.1) we can compute the ratio of volume of this annulus to the volume of the unit ball:

$$\frac{\text{vol}[\mathcal{B}_d(1) \setminus \mathcal{B}_d(1 - \epsilon)]}{\text{vol}(\mathcal{B}_d(1))} = 1 - \epsilon^d.$$

This ratio tends to 1 exponentially fast as $d \rightarrow \infty$. The ratio of volume of the ball $\mathcal{B}_d(1 - \epsilon)$ to the volume of the unit ball $\mathcal{B}_d(1)$ is, similarly to the case of the cubes above, $(1 - \epsilon)^d$. This result extends to any measurable set $A \subset \mathbb{R}^d$. Indeed, define the set $A_{1-\epsilon} = \{(1 - \epsilon)x : x \in A\}$. Then, by splitting A and $A_{1-\epsilon}$ into infinitesimal cubes and adding up their volumes, we find $\text{vol}(A_{1-\epsilon}) = (1 - \epsilon)^d \text{vol}(A)$.

7.6.4 The area of volume concentration in a cube

Let $X = (x_1, \dots, x_d)$ be uniformly distributed on $\mathcal{C}_d = [-1, 1]^d$. Then x_1^2, \dots, x_d^2 are independent r.v. on $[0, 1]$. Hoeffding's inequality gives

$$\mathbb{P}\left\{\left|\frac{1}{d}(x_1^2 + \dots + x_d^2) - \frac{1}{d}\mathbb{E}(x_1^2 + \dots + x_d^2)\right| \geq \epsilon\right\} \leq 2e^{-2d\epsilon^2}.$$

Since $\mathbb{E}x_i^2 = \frac{1}{3}$, we obtain

$$\mathbb{P} \left\{ \left| \|X\|^2 - \frac{d}{3} \right| \geq \epsilon d \right\} \leq 2e^{-2d\epsilon^2}.$$

Therefore, the main volume in the cube \mathcal{C}_d is concentrated in the annulus around the sphere with radius $\sqrt{d/3}$.

7.6.5 Squared norm of a random point in a cube

Let $Z = (z_1, \dots, z_d)$ be a random vectors on $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$ consisting of i.i.d. random components z_i having a distribution with density $p(t)$, $t \in [-\delta, \delta]$, $\delta > 0$.

Set $\eta = \sum_{j=1}^d z_j^2$. We have $\mathbb{E}\eta = d\mu_2$ and $\text{var}(\eta) = d\text{var}(z_1^2) = d(\mu_4 - \mu_2^2)$, where μ_j be the moments of the distribution with density $p(t)$.

For example, when z_i have Beta(α, α) distribution with density

$$p_{\alpha, \delta}(t) = \frac{(2\delta)^{1-2\alpha}}{\text{Beta}(\alpha, \alpha)} [\delta^2 - t^2]^{\alpha-1}, \quad -\delta < t < \delta, \alpha > 0, \quad (7.6.3)$$

where Beta(\cdot, \cdot) is the Beta-function, then

$$\mu_2 = \frac{\delta^2}{2\alpha + 1}, \quad \mu_4 = \frac{3\delta^4}{(2\alpha + 1)(2\alpha + 3)} \quad (7.6.4)$$

and therefore

$$\mathbb{E}\eta = \frac{d\delta^2}{2\alpha + 1}, \quad \text{var}(\eta) = \frac{4d\delta^4\alpha}{(2\alpha + 1)^2(2\alpha + 3)}. \quad (7.6.5)$$

If $\alpha = 1$, when Z is uniform in the cube $\mathcal{C}_d(\delta)$, then

$$\mathbb{E}\eta = \frac{1}{3}d\delta^2, \quad \text{var}(\theta) = \frac{4}{45}d\delta^4. \quad (7.6.6)$$

7.6.6 Distance between two random points in a cube

Assume $Z = (z_1, \dots, z_d)$ and $Z' = (z'_1, \dots, z'_d)$ are independent random vectors on $\mathcal{C}_d(\delta) = [-\delta, \delta]^d$ consisting of i.i.d. random components z_i and z'_i which have some distribution with density $p(t)$, $t \in [-\delta, \delta]$, $\delta > 0$. Let μ_j be the moments of the distribution with density $p(t)$. Assume that the density $p(t)$ is symmetric around 0 and hence all odd moments are zero: $\mu_{2k+1} = 0$ for $k = 1, 2, \dots$

The distribution of the squared distances

$$\theta = \|Z - Z'\|^2 = \sum_{i=1}^d (z_i - z'_i)^2$$

has the mean and variance that can be easily computed as follows:

$$\begin{aligned} \mathbb{E}\theta &= d\mathbb{E}(z_1 - z'_1)^2 = 2d\mu_2, \\ \text{var}(\theta) &= d\text{var}(z_1 - z'_1)^2 = d \left[\mathbb{E}(z_1 - z'_1)^4 - [\mathbb{E}(z_1 - z'_1)^2]^2 \right] = 2d [\mu_4 + \mu_2^2] \end{aligned}$$

For example, when z_i and z'_i have Beta(α, α) distribution with density (7.6.3) and hence moments (7.6.4), we obtain

$$\mathbb{E}\theta = \frac{2d\delta^2}{2\alpha + 1}, \quad \text{var}(\theta) = \frac{4d\delta^4(4\alpha + 3)}{(2\alpha + 1)^2(2\alpha + 3)}. \quad (7.6.7)$$

If $\alpha = 1$ (that is, when Z and Z' are uniform in the cube $\mathcal{C}_d(\delta)$), then

$$\mathbb{E}\theta = \frac{2}{3}d\delta^2, \quad \text{var}(\theta) = \frac{28}{45}d\delta^4 \quad (7.6.8)$$

7.6.7 Volume of the intersection of two balls of the same radius

Let $\mathcal{B}_d(Z_j, r)$ and $\mathcal{B}_d(Z_i, r)$ be two balls in \mathbb{R}^d with same radius and different centers Z and Z' . To compute the volume of the intersection $\mathcal{B}_d(Z, r) \cap \mathcal{B}_d(Z', r)$, we will use the formula, see , for the volume of the d -dimensional cap (cut in the direction of Z') of height h from a d -dimensional ball $\mathcal{B}_d(Z, r)$:

$$K_{d,r,h} = \frac{1}{2}r^d V_d I_{1-h^2/r^2} \left(\frac{d-1}{2}, \frac{1}{2} \right) - \frac{h}{d}(r^2 - h^2)^{(d-1)/2} V_{d-1}, \quad (7.6.9)$$

where V_d is defined in (7.6.1), $\Gamma(\cdot)$ is the Gamma-function and

$$I_t(\alpha, \beta) = \int_0^t u^{\alpha-1}(1-u)^{\beta-1} du \Big/ \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du$$

is the normalised incomplete Beta-function. In the rhs of (7.6.9), the first term is the volume of the related d -dimensional hyper-sector (this expression is derived in [65]) and the second term is the volume of the cone with height h and base $\mathcal{B}_{d-1}((Z + Z')/2, r')$, where $r' = \sqrt{r^2 - h^2}$.

The volume of the intersection of the balls $\mathcal{B}_d(Z, r)$ and $\mathcal{B}_d(Z', r)$ is therefore

$$\text{vol}(\mathcal{B}_d(Z, r) \cap \mathcal{B}_d(Z', r)) = 2K_{d,r,h} \quad (7.6.10)$$

where $h = \frac{1}{2}\|Z - Z'\|$ and $K_{d,r,h}$ is defined in (7.6.9).

7.6.8 A direct computation of $C_{d,Z,r}$

For computing values of $C_{d,Z,r}$, we can employ the following direct approach based on the use of characteristic functions (c.f.).

- (a) Compute the c.f. $\psi_z(s) = \int e^{its} \varphi_z(t) dt$ for $z = z_j$ ($j = 1, \dots, d$), with the density $\varphi_z(t)$ defined either by (7.2.5) or (7.2.7).
- (b) As u_j are independent, the c.f. of $\|U - Z\|^2$ is the product $\psi_Z(s) = \prod_{j=1}^d \psi_{z_j}(s)$.
- (c) The density of $\|U - Z\|^2$ is found using the inversion formula

$$p_{d,Z}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} \psi_Z(s) ds, \quad x \geq 0.$$

For computing the c.f. $\psi_z(s) = \int e^{its} \varphi_z(t) dt$ we can use the formula

$$\int_a^b \frac{e^{xt}}{\sqrt{t}} dt = 2 \int_{\sqrt{a}}^{\sqrt{b}} e^{xu^2} du = \sqrt{\frac{\pi}{x}} \left(\operatorname{erfi}(\sqrt{bx}) - \operatorname{erfi}(\sqrt{ax}) \right)$$

for any $0 \leq a < b < \infty$ and any complex $x \neq 0$. Here $\operatorname{erfi}(x)$ is the imaginary error function

$$\operatorname{erfi}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2} dt = \frac{2}{\sqrt{\pi}} \sum_{j=0}^{\infty} \frac{x^{2j+1}}{j!(2j+1)};$$

the series in the right-hand side of this formula converges for all complex x .

This approach allows very accurate computation of $C_{d,Z,r}$ but it is very computationally intensive and can only be performed for given Z .

7.7 Appendix B: Important auxiliary results

Lemma 1. *Let $\delta > 0$, $x \in \mathbb{R}$ and $\eta_{x,\delta}$ be a r.v. $\eta_{x,\delta} = (\xi - x)^2$, where r.v. ξ has uniform distribution on $[-\delta, \delta]$. Then the c.d.f. of the r.v. $\eta_{x,\delta}$ is*

$$F_{x,\delta}(t) = \mathbb{P}\{\eta_{x,\delta} \leq t\} = \begin{cases} 0 & \text{for } t \leq 0 \\ \frac{\sqrt{t}}{\delta} \cdot 1_{[|x| \leq \delta]} & \text{for } 0 < t < (\delta - |x|)^2 \\ \frac{\delta - |x| + \sqrt{t}}{2\delta} & \text{for } (\delta - |x|)^2 \leq t \leq (\delta + |x|)^2 \\ 1 & (\delta + |x|)^2 < t, \end{cases} \quad (7.7.1)$$

where

$$1_{[|x| \leq \delta]} = \begin{cases} 1 & \text{if } |x| \leq \delta \\ 0 & \text{if } |x| > \delta. \end{cases}$$

The corresponding density of $\eta_{x,\delta}$ is

$$\varphi_{x,\delta}(t) = \begin{cases} 1/(2\delta\sqrt{t}) \cdot 1_{[|x| \leq \delta]} & \text{for } 0 < t < (\delta - |x|)^2 \\ 1/(4\delta\sqrt{t}) & \text{for } (\delta - |x|)^2 < t \leq (\delta + |x|)^2 \\ 0 & \text{otherwise.} \end{cases} \quad (7.7.2)$$

The first four central moments of the r.v. $\eta_{x,\delta}$ are:

$$\mu_{x,\delta}^{(1)} = E\eta_{x,\delta} = x^2 + \frac{\delta^2}{3}, \quad \mu_{x,\delta}^{(2)} = \operatorname{var}(\eta_{x,\delta}) = \frac{4\delta^2}{3} \left(x^2 + \frac{\delta^2}{15} \right), \quad (7.7.3)$$

$$\mu_{x,\delta}^{(3)} = E[\eta_{x,\delta} - E\eta_{x,\delta}]^3 = \frac{16\delta^4}{15} \left(x^2 + \frac{\delta^2}{63} \right), \quad (7.7.4)$$

$$\mu_{x,\delta}^{(4)} = E[\eta_{x,\delta} - E\eta_{x,\delta}]^4 = 3\mu_{x,\delta}^{(1)}\mu_{x,\delta}^{(3)}. \quad (7.7.5)$$

Proof. Clearly, if $t \leq 0$ then $F_{x,\delta}(t) = 0$ and so we only consider the case $t > 0$. In view of symmetry, for all $x \in \mathbb{R}$, $\delta > 0$ and $t \geq 0$, we have $F_{x,\delta}(t) = F_{-x,\delta}(t)$ and therefore we only need to consider $x \geq 0$. Also, $\eta_{x,\delta} \leq (|x| + \delta)^2$ with probability 1 implying $F_{x,\delta}(t) = 1$ for all $t \geq (|x| + \delta)^2$.

Assume $0 \leq x \leq \delta$. We then have for all $t \geq 0$:

$$\begin{aligned} F_{x,\delta}(t) &= \mathbb{P}\{(\xi - x)^2 \leq t\} = \mathbb{P}\{(\xi - x)^2 \leq t, \xi \leq x\} + \mathbb{P}\{(\xi - x)^2 \leq t, \xi > x\} \\ &= \mathbb{P}\{x - \xi \leq \sqrt{t}, \xi \leq x\} + \mathbb{P}\{\xi - x \leq \sqrt{t}, \xi > x\} \\ &= \mathbb{P}\{x - \sqrt{t} \leq \xi \leq x\} + \mathbb{P}\{x < \xi \leq x + \sqrt{t}\} \end{aligned}$$

with

$$\mathbb{P}\{x - \sqrt{t} \leq \xi \leq x\} = \begin{cases} \sqrt{t}/(2\delta) & \text{if } \sqrt{t} < x + \delta \\ (x + \delta)/(2\delta) & \text{if } \sqrt{t} \geq x + \delta, \end{cases}$$

$$\mathbb{P}\{x < \xi \leq x + \sqrt{t}\} = \begin{cases} \sqrt{t}/(2\delta) & \text{if } \sqrt{t} < \delta - x \\ (\delta - x)/(2\delta) & \text{if } \sqrt{t} \geq \delta - x. \end{cases}$$

This yields the expression (7.7.1) for $F_{x,\delta}(t)$ in the case $|x| \leq \delta$.

If $x > \delta$ then $\eta_{x,\delta} \geq (x - \delta)^2$ with probability 1 implying $F_{x,\delta}(t) = 0$ for all $t \leq (x - \delta)^2$ and $\mathbb{P}\{x < \xi \leq x + \sqrt{t}\} = 0$ for all t . Therefore

$$F_{x,\delta}(t) = \mathbb{P}\{x - \sqrt{t} \leq \xi \leq x\} = \begin{cases} 0 & \text{if } \sqrt{t} \leq x - \delta \\ \frac{\delta - (x - \sqrt{t})}{2\delta} & \text{if } x - \delta < \sqrt{t} < x + \delta \\ 1 & \text{if } \sqrt{t} \geq x + \delta, \end{cases}$$

This yields the expression (7.7.1) for $F_{x,\delta}(t)$ in the case $|x| > \delta$.

Deduction of the formulas (7.7.2) for the density and (7.7.3) for the moments from the expression (7.7.1) for the c.d.f. $F_{x,\delta}(t)$ is an easy exercise. □

Lemma 2. *Let $\delta > 0$, $x \in \mathbb{R}$ and $\eta'_{x,\delta}$ be a r.v. $\eta'_{x,\delta} = |\xi - x|$, where r.v. ξ has uniform distribution on $[-\delta, \delta]$. Then the c.d.f. of the r.v. $\eta'_{x,\delta}$ is*

$$F'_{x,\delta}(t) = \mathbb{P}\{\eta'_{x,\delta} \leq t\} = \begin{cases} 0 & \text{for } t \leq 0 \\ \frac{t}{\delta} \cdot \mathbf{1}_{[|x| \leq \delta]} & \text{for } 0 < t < |\delta - |x|| \\ \frac{\delta - |x| + t}{2\delta} & \text{for } |\delta - |x|| \leq t \leq \delta + |x| \\ 1 & \delta + |x| < t, \end{cases} \quad (7.7.6)$$

The corresponding density of $\eta'_{x,\delta}$ is

$$\varphi'_{x,\delta}(t) = \begin{cases} \frac{1}{\delta} \cdot \mathbf{1}_{[|x| \leq \delta]} & \text{for } 0 < t < |\delta - |x|| \\ \frac{1}{2\delta} & \text{for } |\delta - |x|| < t \leq \delta + |x| \\ 0 & \text{otherwise.} \end{cases} \quad (7.7.7)$$

Lemma 2 follows from Lemma 1 by noting that $\eta'_{x,\delta} = \sqrt{\eta_{x,\delta}}$.

Note that $\mathbf{1}_{[|x| \leq \delta]} = 0$ for $|x| > \delta$ and one of the two non-trivial cases in (7.7.1), (7.7.2), (7.7.6) and (7.7.7), when $|x| > \delta$, become trivial as expressions vanish to zero.

Chapter 8

Efficient quantization and weak covering of high dimensional cubes

Abstract

Let $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ be a design; that is, a collection of n points $Z_j \in [-1, 1]^d$. In this chapter, we study the quality of quantization of $[-1, 1]^d$ by the points of \mathbb{Z}_n and the problem of quality of covering of $[-1, 1]^d$ by $\mathcal{B}_d(\mathbb{Z}_n, r)$, the union of balls centred at $Z_j \in \mathbb{Z}_n$. We concentrate on the cases where the dimension d is not small ($d \geq 5$) and n is not too large, $n \leq 2^d$. As a result of the strong performance of Scheme 3 in Chapter 7, we define the design $\mathbb{D}_{n,\delta}$ as the maximum-resolution 2^{d-1} design defined on vertices of the cube $[-\delta, \delta]^d$, $0 \leq \delta \leq 1$. For this design, we derive a closed-form expression for the quantization error and very accurate approximations for the coverage area $\text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$. It is conjectured that the design $\mathbb{D}_{n,\delta}$ with optimal δ is the most efficient quantizer of $[-1, 1]^d$ under the assumption $n \leq 2^d$ and it is also makes a very efficient $(1-\gamma)$ -covering. The results of a large-scale numerical investigation confirming the accuracy of the developed approximations and the efficiency of the designs $\mathbb{D}_{n,\delta}$ is provided. The content of this chapter has been submitted for publication and can be viewed also at [85].

8.1 Introduction

8.1.1 Main notation

- $\|\cdot\|$: the Euclidean norm;
- $\mathcal{B}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| \leq r\}$: d -dimensional ball of radius r centered at $Z \in \mathbb{R}^d$;
- $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$: a design; that is, a collection of n points $Z_j \in \mathbb{R}^d$;
- $\mathcal{B}_d(\mathbb{Z}_n, r) = \bigcup_{j=1}^n \mathcal{B}_d(Z_j, r)$;
- $C_d(\mathbb{Z}_n, r) = \text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))/2^d$: the proportion of the cube $[-1, 1]^d$ covered by $\mathcal{B}_d(\mathbb{Z}_n, r)$;
- vectors in \mathbb{R}^d are row-vectors;

- for any $a \in \mathbb{R}$, $\mathbf{a} = (a, a, \dots, a) \in \mathbb{R}^d$.

8.1.2 Main problems of interest

We will study the following two main characteristics of designs \mathbb{Z}_n .

1. Quantization error. Let $X = (x_1, \dots, x_d)$ be uniform random vector on $[-1, 1]^d$. The mean squared quantization error for a design $\mathbb{Z}_n = \{Z_1, \dots, Z_n\} \subset \mathbb{R}^d$ is defined by

$$\theta(\mathbb{Z}_n) = \mathbb{E}_X \varrho^2(X, \mathbb{Z}_n), \quad \text{where } \varrho^2(X, \mathbb{Z}_n) = \min_{Z_i \in \mathbb{Z}_n} \|X - Z_i\|^2. \quad (8.1.1)$$

2. Weak coverage. Denote the proportion of the cube $[-1, 1]^d$ covered by the union of n balls $\mathcal{B}_d(\mathbb{Z}_n, r) = \bigcup_{j=1}^n \mathcal{B}_d(Z_j, r)$ by

$$C_d(\mathbb{Z}_n, r) := \text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r)) / 2^d.$$

For given radius $r > 0$, the union of n balls $\mathcal{B}_d(\mathbb{Z}_n, r)$ makes the $(1 - \gamma)$ -coverage of the cube $[-1, 1]^d$ if

$$C_d(\mathbb{Z}_n, r) = 1 - \gamma. \quad (8.1.2)$$

Complete coverage corresponds to $\gamma = 0$. In this chapter, the complete coverage of $[-1, 1]^d$ will not be enforced and we will mostly be interested in *weak coverage*, that is, achieving (8.1.2) with some small $\gamma > 0$.

Two n -point designs \mathbb{Z}_n and \mathbb{Z}'_n will be differentiated in terms of performance as follows: (a) \mathbb{Z}_n dominates \mathbb{Z}'_n for quantization if $\theta(\mathbb{Z}_n) < \theta(\mathbb{Z}'_n)$; (b) if for a given $\gamma \geq 0$, $C_d(\mathbb{Z}_n, r_1) = C_d(\mathbb{Z}'_n, r_2) = 1 - \gamma$ and $r_1 < r_2$, then the design \mathbb{Z}_n provides a more efficient $(1 - \gamma)$ -covering than \mathbb{Z}'_n and is therefore preferable. In Section 8.1.5 we extend these definitions by allowing the two designs to have different number of points and, moreover, to have different dimensions.

8.1.3 Relation between quantization and weak coverage

The two characteristics, $C_d(\mathbb{Z}_n, r)$ and $\theta(\mathbb{Z}_n)$, are related: $C_d(\mathbb{Z}_n, r)$, as a function of $r \geq 0$, is the c.d.f. of the r.v. $\varrho(X, \mathbb{Z}_n)$ while $\theta(\mathbb{Z}_n)$ is the second moment of the distribution with this c.d.f.:

$$\theta(\mathbb{Z}_n) = \int_{r \geq 0} r^2 dC_d(\mathbb{Z}_n, r). \quad (8.1.3)$$

In particular, this yields that if an n -point design \mathbb{Z}_n^* maximizes, in the set of all n -point designs, $C_d(\mathbb{Z}_n, r)$ for all $r > 0$, then it also minimizes $\theta(\mathbb{Z}_n)$. Moreover, if r.v. $\varrho(X, \mathbb{Z}_n)$ stochastically dominates $\varrho(X, \mathbb{Z}'_n)$, so that $C_d(\mathbb{Z}'_n, r) \leq C_d(\mathbb{Z}_n, r)$ for all $r \geq 0$ and the inequality is strict for at least one r , then $\theta(\mathbb{Z}_n) < \theta(\mathbb{Z}'_n)$.

The relation (8.1.3) can alternatively be written as

$$\theta(\mathbb{Z}_n) = \int_{r \geq 0} r dC_d(\mathbb{Z}_n, \sqrt{r}), \quad (8.1.4)$$

where $C_d(\mathbb{Z}_n, \sqrt{r})$, considered as a function of r , is the c.d.f. of the r.v. $\varrho^2(X, \mathbb{Z}_n)$ and hence $\theta(\mathbb{Z}_n)$ is the mean of this r.v. Relation (8.1.4) is simply another form of (8.1.1).

8.1.4 Re-normalization of the quantization error

To compare efficiency of n -point designs \mathbb{Z}_n with different values of n , one must suitably normalise $\theta(\mathbb{Z}_n)$ with respect to n . A classical characteristic for quantization in space, as formulated in [19, f-la (86), Ch.2], we obtain

$$Q_d(\mathbb{Z}_n) := \frac{1}{d} \frac{\frac{1}{n} \sum_{i=1}^n \int_{V(Z_i)} \|X - Z_i\|^2 dX}{\left[\frac{1}{n} \sum_{i=1}^n \text{vol}(V(Z_i))\right]^{1+\frac{2}{d}}}. \quad (8.1.5)$$

Note that $Q_d(\mathbb{Z}_n)$ is re-normalised with respect to dimension d too, not only with respect to n . Normalization $1/d$ with respect to d is very natural in view of the definition of the Euclidean norm.

Using (8.2.1) below, for the cube $[-1, 1]^d$ the quantity in (8.1.5) can be expressed as

$$Q_d(\mathbb{Z}_n) = \frac{n^{2/d} \theta(\mathbb{Z}_n)}{d \left[\sum_{i=1}^n \text{vol}(V(Z_i))\right]^{2/d}} = \frac{n^{2/d} \theta(\mathbb{Z}_n)}{d \cdot \text{vol}([-1, 1]^d)^{2/d}} = \frac{n^{2/d}}{4d} \theta(\mathbb{Z}_n). \quad (8.1.6)$$

The normalising factor $n^{2/d}$ is explained in Section 7.5 of Chapter 7.

8.1.5 Renormalised versions and formulation of optimal design problems

In view of (8.1.6), the naturally defined re-normalized version of $\theta(\mathbb{Z}_n)$ is $Q_d(\mathbb{Z}_n) = n^{2/d} \theta(\mathbb{Z}_n) / (4d)$. From (8.1.3) and (8.1.4), $Q_d(\mathbb{Z}_n)$ is the expectation of the r.v. $n^{2/d} \varrho^2(X, \mathbb{Z}_n) / (4d)$ and the second moment of the r.v. $n^{1/d} \varrho(X, \mathbb{Z}_n) / (2\sqrt{d})$ respectively. This suggests the following re-normalization of the radius r with respect to n and d :

$$R = n^{1/d} r / (2\sqrt{d}). \quad (8.1.7)$$

We can then define optimal designs as follows. Let d be fixed, $\mathcal{Z}_n = \{\mathbb{Z}_n\}$ be the set of all n -point designs and $\mathcal{Z} = \cup_{n=1}^{\infty} \mathcal{Z}_n$ be the set of all designs.

Definition 8.1.1 *The design \mathbb{Z}_m^* with some m is optimal for quantization in $[-1, 1]^d$, if*

$$Q_d(\mathbb{Z}_m^*) = \min_n \min_{\mathbb{Z}_n \in \mathcal{Z}_n} Q_d(\mathbb{Z}_n) = \min_{\mathbb{Z} \in \mathcal{Z}} Q_d(\mathbb{Z}). \quad (8.1.8)$$

Definition 8.1.2 *The design \mathbb{Z}_m^* with some m is optimal for $(1 - \gamma)$ -coverage of $[-1, 1]^d$, if*

$$R_{1-\gamma}(\mathbb{Z}_m^*) = \min_n \min_{\mathbb{Z}_n \in \mathcal{Z}_n} R_{1-\gamma}(\mathbb{Z}_n) = \min_{\mathbb{Z} \in \mathcal{Z}} R_{1-\gamma}(\mathbb{Z}). \quad (8.1.9)$$

Here $0 \leq \gamma \leq 1$ and for a given design $\mathbb{Z}_n \in \mathcal{Z}_n$,

$$R_{1-\gamma}(\mathbb{Z}_n) = n^{1/d} r_{1-\gamma}(\mathbb{Z}_n) / (2\sqrt{d}), \quad (8.1.10)$$

where $r_{1-\gamma}(\mathbb{Z}_n)$ is defined as the smallest r such that $C_d(\mathbb{Z}_n, r) = 1 - \gamma$.

Importance of the factor \sqrt{d} in (8.1.7) will be seen in Section 8.3.6 where we shall study the asymptotical behaviour of $(1 - \gamma)$ -coverings for large d .

8.1.6 Thickness of covering

Let $\gamma = 0$ in Definition 8.1.2. Then $r_1(\mathbb{Z}_n)$ is the covering radius associated with \mathbb{Z}_n so that the union of the balls $\mathcal{B}_d(\mathbb{Z}_n, r)$ with $r = r_1(\mathbb{Z}_n)$ makes a covering of $[-1, 1]^d$. Let us tile up the whole space \mathbb{R}^d with the translations of the cube $[-1, 1]^d$ and corresponding translations of the balls $\mathcal{B}_d(\mathbb{Z}_n, r)$. This would make a full covering of the whole space; denote this space covering by $\mathcal{B}_d(\mathbb{Z}_n, r)$. The thickness Θ of any space covering is defined, see [19, f-la (1), Ch. 2], as the average number of balls containing a point of the whole space. In our case of $\mathcal{B}_d(\mathbb{Z}_n, r)$, the thickness is

$$\Theta(\mathcal{B}_d(\mathbb{Z}_n, r)) = \frac{n \operatorname{vol}(\mathcal{B}_d(0, r))}{\operatorname{vol}([-1, 1]^d)} = \frac{n r^d \operatorname{vol}(\mathcal{B}_d(0, 1))}{2^d}.$$

The normalised thickness, θ , is the thickness Θ divided by $\operatorname{vol}(\mathcal{B}_d(0, 1))$, the volume of the unit ball, see [19, f-la (2), Ch. 2]. In the case of $\mathcal{B}_d(\mathbb{Z}_n, r)$, the normalised thickness is

$$\theta(\mathcal{B}_d(\mathbb{Z}_n, r)) = \frac{n r^d}{2^d} = d^{d/2} [R_1(\mathbb{Z}_n)]^d,$$

where we have recalled that $r = r_1(\mathbb{Z}_n)$ and $R_{1-\gamma}(\mathbb{Z}_n) = n^{1/d} r_{1-\gamma}(\mathbb{Z}_n) / (2\sqrt{d})$ for any $0 \leq \gamma \leq 1$.

We can define the normalised thickness of the covering of the cube by the same formula and extend it to any $0 \leq \gamma \leq 1$:

Definition 8.1.3 *Let $\mathcal{B}_d(\mathbb{Z}_n, r)$ be a $(1 - \gamma)$ -covering of the cube $[-1, 1]^d$ with $0 \leq \gamma \leq 1$. Its normalised thickness is defined by*

$$\theta(\mathcal{B}_d(\mathbb{Z}_n, r)) = (\sqrt{d}R)^d, \tag{8.1.11}$$

where $R = n^{1/d} r / (2\sqrt{d})$, see (8.1.7).

In view of (8.1.11), we can reformulate the definition (8.1.9) of the $(1 - \gamma)$ -covering optimal design by saying that this design minimizes (normalised) thickness in the set of all $(1 - \gamma)$ -covering designs.

8.1.7 The design of the main interest

We will be mostly interested in the following n -point design $\mathbb{Z}_n = \mathbb{D}_{n,\delta}$ defined only for $n = 2^{d-1}$:

Design $\mathbb{D}_{n,\delta}$: *a maximum-resolution 2^{d-1} design defined on vertices of the cube $[-\delta, \delta]^d$, $0 \leq \delta \leq 1$.*

The design $\mathbb{D}_{n,1/2}$ extends to the lattice D_d (shifted by $\frac{1}{2}$) containing points $X = (x_1, \dots, x_d)$ with integer components satisfying $x_1 + \dots + x_d = 0 \pmod{2}$, see [19, Sect. 7.1, Ch. 4]; this lattice is sometimes called ‘checkerboard lattice’. The motivation to theoretically study the design $\mathbb{D}_{n,\delta}$ is a consequence of numerical results reported for Scheme 3 in Chapter 7 (see also [144]). In Chapter 7 it is shown that for all dimensions $d \geq 7$, the design $\mathbb{D}_{n,\delta}$ with suitable δ provides the best

quantization and covering per point among all other designs considered. In Chapter 7, we consider n -point designs in d -dimensional cubes providing good covering and quantization. Aiming at practical applications in computer experiments, optimization and numerical integration, the aim was to consider the designs with n which is not too large and in any case does not exceed 2^d . In fact, as a result of extensive numerical comparisons performed in Chapter 7 and the analysis performed in this chapter, the author states the following conjecture.

Conjecture. *For all $d \geq 7$, the design $\mathbb{D}_{n,\delta}$ with optimal δ determined by (8.2.11), is the optimal design for quantization, if we add the restriction $m \leq 2^d$ in (8.1.8).*

The author does not think that the conjecture is true without a restriction on m in (8.1.8) as for very large m we can construct one of the very efficient lattice space quantizers, see [19, Sect. 3, Ch. 2], take the lattice points belonging to a very large cube and scale the cube back to $[-1, 1]^d$; this may result in a better quantizer than the design $\mathbb{D}_{n,\delta}$. It is difficult to study (both, numerically and theoretically) properties of such designs since they have to have very large number of points n and are expected to have several non-congruent types of Voronoi cells due to boundary conditions. As we are interested in practical applications, designs with practically reachable values of n are more important. The author does not state a conjecture concerning the structure of the best covering designs as formulated in Definition 2, as the structure of such designs depends on both d and γ , see discussions in Section 8.4.

For theoretical comparison with design $\mathbb{D}_{n,\delta}$, we shall consider the following simple design, which extends to the integer point lattice Z^d (shifted by $\frac{1}{2}$) in the whole space \mathbb{R}^d :

Design $\mathbb{D}_n^{(0)}$: *the collection of 2^d points $(\pm\frac{1}{2}, \dots, \pm\frac{1}{2})$, all vertices of the cube $[-\frac{1}{2}, \frac{1}{2}]^d$.*

Without loss of generality, while considering the design $\mathbb{D}_{n,\delta}$ we assume that the point $Z_1 \in \mathbb{D}_{n,\delta} = \{Z_1, \dots, Z_n\}$ is $Z_1 = \boldsymbol{\delta} = (\delta, \dots, \delta)$. Similarly, the first point in $\mathbb{D}_n^{(0)}$ is $Z_1 = \frac{1}{2} = (\frac{1}{2}, \dots, \frac{1}{2})$. Note also that for numerical comparisons, in Section 8.4 we shall introduce one more design.

8.1.8 Structure of the rest of the chapter and the main results

In Section 8.2 we study $Q_d(\mathbb{D}_{n,\delta})$, the normalized mean squared quantization error for the design $\mathbb{D}_{n,\delta}$. There are two important results, Theorems 8.2.1 and 8.2.2. In Theorems 8.2.1, we derive the explicit form for the Voronoi cells for the points of the design $\mathbb{D}_{n,\delta}$ and in Theorem 8.2.2 we derive a closed-form expression for $Q_d(\mathbb{D}_{n,\delta})$ for any $\delta > 0$. As a consequence, in Corollary 8.2.1 we determine the value of the optimal value of δ .

The main result of Section 8.3 is Theorem 8.3.1, where we derive closed-form expressions (in terms of $C_{d,Z,r}$, the fraction of the cube $[-1, 1]^d$ covered by a ball

$\mathcal{B}_d(Z, r)$) for the coverage area with $\text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$. Then, using accurate approximations for $C_{d,Z,r}$, we derive approximations for $\text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$. In Theorem 8.3.2 we derive asymptotic expressions for the $(1 - \gamma)$ -coverage radius for the design $\mathbb{D}_{n,1/2}$ and show that for any $\gamma > 0$, the ratio of the $(1 - \gamma)$ -coverage radius to the 1-coverage radius tends to $1/\sqrt{3}$ as $d \rightarrow \infty$. Numerical results of Section 8.3.6 confirm that even for rather small d , the 0.999-coverage radius is much smaller than the 1-coverage radius providing the full coverage.

In Section 8.4 we demonstrate that the approximations developed in Section 8.3 are very accurate and make a comparative study of selected designs used for quantization and covering.

In Appendices A–C located at the end of this chapter, proofs of the most technical results are provided.

The main results of this chapter are: a) derivation of the closed-form expression for the quantization error for the design $\mathbb{D}_{n,\delta}$, and b) derivation of accurate approximations for the coverage area $\text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r))$ for the design $\mathbb{D}_{n,\delta}$.

8.2 Quantization

8.2.1 Reformulation in terms of the Voronoi cells

Consider any n -point design $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$. The Voronoi cell $V(Z_i)$ for $Z_i \in \mathbb{Z}_n$ is defined as

$$V(Z_i) = \{x \in [-1, 1]^d : \|Z_i - x\| \leq \|Z_j - x\| \text{ for } j \neq i\}.$$

The mean squared quantization error $\theta(\mathbb{Z}_n)$ introduced in (8.1.1) can be written in terms of the Voronoi cells as follows:

$$\theta(\mathbb{Z}_n) = \mathbb{E}_X \min_{i=1, \dots, n} \|X - Z_i\|^2 = \frac{1}{\text{vol}([-1, 1]^d)} \sum_{i=1}^n \int_{V(Z_i)} \|X - Z_i\|^2 dX, \quad (8.2.1)$$

where $X = (x_1, \dots, x_d)$ and $dX = dx_1 dx_2 \cdots dx_d$.

This reformulation has significant benefit when the design \mathbb{Z}_n has certain regular structure. In particular, if all of the Voronoi cells $V(Z_i), i = 1, \dots, n$, are congruent, then we can simplify (8.2.1) to

$$\theta(\mathbb{Z}_n) = \frac{1}{\text{vol}(V(Z_1))} \int_{V(Z_1)} \|X - Z_1\|^2 dX. \quad (8.2.2)$$

In Section 8.2.3, this formula will be the starting point for derivation of the closed-form expression for $\theta(\mathbb{Z}_n)$ for the design $\mathbb{D}_{d,\delta}$.

8.2.2 Voronoi cells for $\mathbb{D}_{n,\delta}$

Proposition 8.2.1 *Consider the design $\mathbb{D}_{n,\delta}^{(0)}$, the collection of $n = 2^d$ points $(\pm\delta, \dots, \pm\delta)$, $0 < \delta < 1$. The Voronoi cells for this design are all congruent. The Voronoi cell for the point $\delta = (\delta, \delta, \dots, \delta)$ is the cube*

$$C_0 = \left\{ X = (x_1, \dots, x_d) \in \mathbb{R}^d : 0 \leq x_i \leq 1, i = 1, 2, \dots, d \right\}. \quad (8.2.3)$$

Proof. Consider the Voronoi cells created by the design $\mathbb{D}_{n,\delta}^{(0)}$ in the whole space \mathbb{R}^d . For the point $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta)$, the Voronoi cell is clearly $\{X = (x_1, \dots, x_d) : x_i \geq 0\}$. By intersecting this set with the cube $[-1, 1]^d$ we obtain (8.2.3). \square

Theorem 8.2.1 *The Voronoi cells of the design $\mathbb{D}_{n,\delta} = \{Z_1, \dots, Z_n\}$ are all congruent. The Voronoi cell for the point $Z_1 = \boldsymbol{\delta} = (\delta, \delta, \dots, \delta) \in \mathbb{R}^d$ is*

$$V(Z_1) = C_0 \cup \left[\bigcup_{j=1}^d U_j \right] \quad (8.2.4)$$

where C_0 is the cube (8.2.3) and

$$U_j = \left\{ X = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d : -1 \leq x_j \leq 0, |x_j| \leq x_k \leq 1 \text{ for all } k \neq j \right\} \quad (8.2.5)$$

The volume of $V(Z_1)$ is $\text{vol}(V(Z_1)) = 2$.

Proof. The design $\mathbb{D}_{n,\delta}$ is symmetric with respect to all components implying that all $n = 2^{d-1}$ Voronoi cells are congruent immediately yielding that their volumes equal 2. Consider $V(Z_1)$ with $Z_1 = \boldsymbol{\delta}$.

Since $\mathbb{D}_{n,\delta} \subset \mathbb{D}_{n,\delta}^{(0)}$, where design $\mathbb{D}_{n,\delta}^{(0)}$ is introduced in Proposition 8.2.1, and C_0 is the Voronoi set of $\boldsymbol{\delta}$ for design $\mathbb{D}_{n,\delta}^{(0)}$, $C_0 \subset V(\boldsymbol{\delta})$ for design $\mathbb{D}_{n,\delta}$ too.

Consider the d cubes adjacent to C_0 :

$$C_j = \left\{ X = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d : -1 \leq x_j \leq 0, 0 \leq x_i \leq 1 \text{ for all } i \neq j \right\}; \quad j = 1, \dots, d.$$

A part of each cube C_j belongs to $V(Z_1)$. This part is exactly the set U_j defined by (8.2.5). This can be seen as follows. A part of C_j also belongs to the Voronoi set of the point $X_{jk} = \boldsymbol{\delta} - 2\delta e_j - 2\delta e_k$, where $e_l = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 placed at l -th place; all components of X_{jk} are δ except j -th and k -th components which are $-\delta$. We have to have $|x_j| \leq x_k$, for a point $X \in C_j$ to be closer to Z_1 than to X_{jk} . Joining all constraints for $X = (x_1, x_2, \dots, x_d) \in C_j$ ($k = 1, \dots, d, k \neq j$) we obtain (8.2.5) and hence (8.2.4). \square

8.2.3 Explicit formulae for the quantization error

Theorem 8.2.2 *For the design $\mathbb{D}_{n,\delta}$ with $0 \leq \delta \leq 1$, we obtain:*

$$\theta(\mathbb{D}_{n,\delta}) = d \left(\delta^2 - \delta + \frac{1}{3} \right) + \frac{2\delta}{d+1}, \quad (8.2.6)$$

$$Q_d(\mathbb{D}_{n,\delta}) = 2^{-2/d} \left(\delta^2 - \delta + \frac{1}{3} + \frac{2\delta}{d(d+1)} \right). \quad (8.2.7)$$

Proof. To compute $\theta(\mathbb{D}_{n,\delta})$, we use (8.2.2), where, in view of Theorem 8.2.1, $\text{vol}(V(Z_1)) = 2$. Using the expression (8.2.4) for $V(Z_1)$ with $Z_1 = \boldsymbol{\delta}$, we obtain

$$\theta(\mathbb{D}_{n,\delta}) = \frac{1}{2} \int_{V(Z_1)} \|X - Z_1\|^2 dX = \frac{1}{2} \left[\int_{C_0} \|X - Z_1\|^2 dX + d \int_{U_1} \|X - Z_1\|^2 dX \right]. \quad (8.2.8)$$

Consider the two terms in (8.2.8) separately. The first term is easy:

$$\begin{aligned} \int_{C_0} \|X - Z_1\|^2 dX &= \int_{C_0} \sum_{i=1}^d (x_i - \delta)^2 dx_1 \dots dx_d = d \int_0^1 (x - \delta)^2 dx \\ &= d \left(\delta^2 - \delta + \frac{1}{3} \right). \end{aligned} \quad (8.2.9)$$

For the second term we have:

$$\begin{aligned} \int_{U_1} \|X - Z_1\|^2 dX &= \int_{-1}^0 \left[\int_{|x_1|}^1 \dots \int_{|x_1|}^1 \sum_{i=1}^d (x_i - \delta)^2 dx_2 \dots dx_d \right] dx_1 \\ &= \int_{-1}^0 (x_1 - \delta)^2 (1 + x_1)^{d-1} dx_1 \\ &+ (d-1) \int_{-1}^0 (1 + x_1)^{d-2} \int_{|x_1|}^1 (x_2 - \delta)^2 dx_2 dx_1 \\ &= \delta^2 - \delta + \frac{1}{3} + \frac{4\delta}{d(d+1)}. \end{aligned} \quad (8.2.10)$$

Inserting the obtained expressions into (8.2.8) we obtain (8.2.6). The expression (8.2.7) is a consequence of (8.1.6), (8.2.6) and $n = 2^{d-1}$. \square

A simple consequence of Theorem 8.2.2 is the following corollary.

Corollary 8.2.1 *The optimal value of δ minimising $\theta(\mathbb{D}_{n,\delta})$ and $Q_d(\mathbb{D}_{n,\delta})$ is*

$$\delta^* = \frac{1}{2} - \frac{1}{d(d+1)}; \quad (8.2.11)$$

for this value,

$$Q_d(\mathbb{D}_{n,\delta^*}) = \min_{\delta} Q_d(\mathbb{D}_{n,\delta}) = 2^{-2/d} \left[\frac{1}{12} + \frac{d^2 + d - 1}{(d+1)^2 d^2} \right]. \quad (8.2.12)$$

From (8.2.7), for the design $\mathbb{D}_{n,\delta}$ with $\delta = 1/2$ we get

$$Q_d(\mathbb{D}_{n,1/2}) = 2^{-2/d} \left[\frac{1}{12} + \frac{1}{(d+1)d} \right], \quad (8.2.13)$$

which is always slightly larger than (8.2.12). Let us make five more remarks.

1. For the one-point design $\mathbb{D}^{(0)} = \{0\}$ with the single point 0 and the design $\mathbb{D}_n^{(0)}$ with $n = 2^d$ points $(\pm\frac{1}{2}, \dots, \pm\frac{1}{2})$ we have $Q_d(\mathbb{D}^{(0)}) = Q_d(\mathbb{D}_n^{(0)}) = 1/12$, which coincides with the value of Q_d in the case of space quantization by the integer-point lattice Z^d , see [19, Ch. 2 and 21].
2. The quantization error (8.2.13) for the design $\mathbb{D}_{n,1/2}$ have almost exactly the same form as the quantization error for the ‘checkerboard lattice’ D_d in \mathbb{R}^d ; the difference is in the factor 1/2 in the last term in (8.2.13), see [19, f-la (27), Ch.21]. To understand quantization error for a lattice, we consider the quantization of a uniform distribution over a large ball in \mathbb{R}^d , see [19, p.59, Ch.21] for a detailed explanation. Naturally, the quantization error Q_d for D_d in \mathbb{R}^d is slightly smaller than Q_d for $\mathbb{D}_{n,1/2}$ in $[-1, 1]^d$.

3. The optimal value of δ in (8.2.11) is smaller than $1/2$. This is caused by a non-symmetrical shape of the Voronoi cells $V(Z_j)$ for designs $\mathbb{D}_{n,\delta}$, which is clearly visible in (8.2.4).
4. The minimal value of $Q_d(\mathbb{D}_{n,\delta^*})$ is achieved with $d = 15$.
5. Formulas (8.2.11) and (8.2.12) are in agreement with numerical results seen in Table 7.4 of Chapter 7.

Let us briefly illustrate the results above. In Figure 8.1, the black circles depict the quantity $Q_d(\mathbb{D}_{n,\delta^*})$ as a function of d . The quantity $Q_d(\mathbb{D}_n^{(0)}) = 1/12$ is shown with the solid red line. We conclude that from dimension seven onwards, the design \mathbb{D}_{n,δ^*} provides better quantization per points than the design $\mathbb{D}_n^{(0)}$. Moreover for $d > 15$, the quantity $Q_d(\mathbb{D}_{n,\delta^*})$ slowly increases and converges to $1/12$. Typical behaviour of $Q_d(\mathbb{D}_{n,\delta})$ as a function of δ is shown in Figure 8.2. This figure demonstrates the significance of choosing δ optimally.

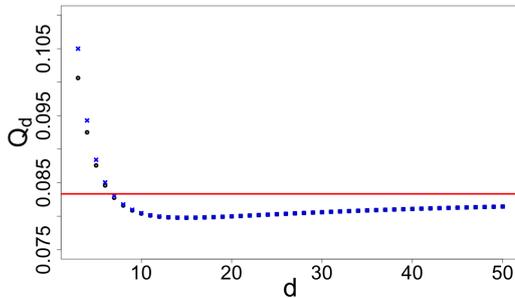


Figure 8.1: $Q_d(\mathbb{D}_{n,\delta^*})$ and $Q_d(\mathbb{D}_{n,1/2})$ as functions of d and $Q_d(\mathbb{D}_n^{(0)}) = 1/12$; $d = 3, \dots, 50$.

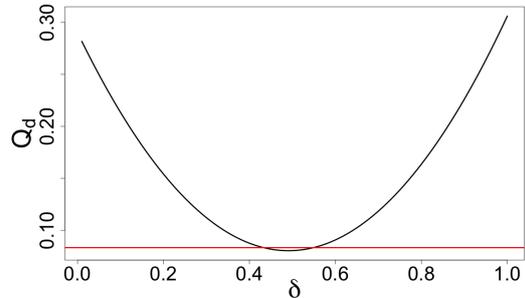


Figure 8.2: $Q_d(\mathbb{D}_{n,\delta})$ as a function of δ and $Q_d(\mathbb{D}_n^{(0)}) = 1/12$ with $d = 10$.

8.3 Closed-form expressions for the coverage area with $\mathbb{D}_{n,\delta}$ and approximations

In this section, we will derive explicit expressions for the coverage area of the cube $[-1, 1]^d$ by the union of the balls $\mathcal{B}_d(\mathbb{D}_{n,\delta}, r)$ associated with the design $\mathbb{D}_{n,\delta}$ introduced in Section 8.1.2. That is, we will derive expressions for the quantity $C_d(\mathbb{D}_{n,\delta}, r)$ for all values of r . Then, in Section 8.3.4, we shall obtain approximations for $C_d(\mathbb{D}_{n,\delta}, r)$. The accuracy of the approximations will be assessed in Section 8.4.2.

8.3.1 Reduction to Voronoi cells

For an n -point design $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$, denote the proportion of the Voronoi cell around Z_i covered by the ball $\mathcal{B}_d(Z_i, r)$ as:

$$V_{d,Z_i,r} := \text{vol}(V(Z_i) \cap \mathcal{B}_d(Z_i, r)) / \text{vol}(V(Z_i)).$$

Then we can state the following simple lemma.

Lemma 8.3.1 Consider a design $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ such that all Voronoi cells $V(Z_i)$ are congruent. Then for any $Z_i \in \mathbb{Z}_n$, $C_d(\mathbb{Z}_n, r) = V_{d, Z_i, r}$.

In view of Theorem 8.2.1, for design $\mathbb{D}_{n, \delta}$ all Voronoi cells $V(Z_i)$ are congruent and $\text{vol}(V(Z_i)) = 2$; recall that $n = 2^{d-1}$. By then applying Lemma 8.3.1 and without loss of generality we have choosen $Z_1 = \boldsymbol{\delta} = (\delta, \delta, \dots, \delta) \in \mathbb{R}^d$, we have for any $r > 0$

$$V_{d, \boldsymbol{\delta}, r} = \frac{1}{2} \text{vol}(V(\boldsymbol{\delta}) \cap \mathcal{B}_d(\boldsymbol{\delta}, r)) = C_d(\mathbb{D}_{n, \delta}, r). \quad (8.3.1)$$

In order to formulate explicit expressions for $V_{d, \boldsymbol{\delta}, r}$, we need an important quantity, proportion of intersection of $[-1, 1]^d$ with one ball.

8.3.2 Intersection of $[-1, 1]^d$ with one ball

Take the cube $[-1, 1]^d$ and a ball $\mathcal{B}_d(Z, r) = \{Y \in \mathbb{R}^d : \|Y - Z\| \leq r\}$ centered at a point $Z = (z_1, \dots, z_d) \in \mathbb{R}^d$; this point Z could be outside $[-1, 1]^d$. The fraction of the cube $[-1, 1]^d$ covered by the ball $\mathcal{B}_d(Z, r)$ is denoted by

$$C_{d, Z, r} = \text{vol}([-1, 1]^d \cap \mathcal{B}_d(Z, r)) / 2^d.$$

The following relation will prove useful. Assume that we have the cube $[-\beta, \beta]^d$ of volume $(2\beta)^d$, the ball $\mathcal{B}_d(Z', r') = \{Y \in \mathbb{R}^d : \|Y - Z'\| \leq r'\}$ with a center at a point $Z' = (z'_1, \dots, z'_d)$. Denote the fraction of the cube $[-\beta, \beta]^d$ covered by the ball $\mathcal{B}_d(Z', r')$ by

$$C_{d, Z', r'}^{(\beta)} = \text{vol}([- \beta, \beta]^d \cap \mathcal{B}_d(Z', r')) / (2\beta)^d.$$

Then the change of the coordinates $Z = Z'/\beta = (z'_1/\beta, \dots, z'_d/\beta)$ and the radius $r = r'/\beta$ gives

$$C_{d, Z', r'}^{(\beta)} = C_{d, Z, r}. \quad (8.3.2)$$

8.3.3 Expressing $C_d(\mathbb{D}_{n, \delta}, r)$ through $C_{d, Z, r}$

Theorem 8.3.1 The quantity $C_d(\mathbb{D}_{n, \delta}, r)$ can be expressed through $C_{d, Z, r}$ for suitable Z as follows.

- For $r \leq \delta$:

$$C_d(\mathbb{D}_{n, \delta}, r) = \frac{1}{2} C_{d, 2\boldsymbol{\delta} - \mathbf{1}, 2r}. \quad (8.3.3)$$

- For $\delta \leq r \leq 1 + \delta$:

$$C_d(\mathbb{D}_{n, \delta}, r) = \frac{1}{2} \left[C_{d, 2\boldsymbol{\delta} - \mathbf{1}, 2r} + d \int_0^{r-\delta} C_{d-1, \frac{2\delta-1-x}{1-x}, \frac{2\sqrt{r^2-(x+\delta)^2}}{1-x}} (1-x)^{d-1} dx \right]. \quad (8.3.4)$$

- For $r \geq 1 + \delta$:

$$C_d(\mathbb{D}_{n,\delta}, r) = \frac{1}{2} \left[C_{d,2\delta-1,2r} + d \int_0^1 C_{d-1, \frac{2\delta-1-x}{1-x}, \frac{2\sqrt{r^2-(x+\delta)^2}}{1-x}} (1-x)^{d-1} dx \right]. \quad (8.3.5)$$

The proof of Theorem 8.3.1 is given in Appendix A, see Section 8.5.

8.3.4 Approximation for $C_d(\mathbb{D}_{n,\delta}, r)$

Accurate approximations for $C_{d,Z,r}$ for arbitrary d, Z and r were developed in Chapter 7. By using the general expansion in the central limit theorem for sums of independent non-identical r.v., we obtain from (7.2.14) in Chapter 7:

$$C_{d,Z,r} \cong \Phi(t) + \frac{\|Z\|^2 + d/63}{5\sqrt{3}(\|Z\|^2 + d/15)^{3/2}} (1-t^2)\varphi(t), \quad (8.3.6)$$

where

$$t = \frac{\sqrt{3}(r^2 - \|Z\|^2 - d/3)}{2\sqrt{\|Z\|^2 + d/15}}.$$

Using (8.3.6), we formulate the following approximation for $C_d(\mathbb{D}_{n,\delta}, r)$.

Approximations for $C_d(\mathbb{D}_{n,\delta}, r)$. Approximate the values C_{\dots} in formulas (8.3.3), (8.3.4), (8.3.5) with corresponding approximations (8.3.6).

8.3.5 Simple bounds for $C_d(\mathbb{D}_{n,\delta}, r)$

Lemma 8.3.2 For any $r \geq 0$, $0 < \delta < 1$ and $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta) \in \mathbb{R}^d$, the quantity $C_d(\mathbb{D}_{n,\delta}, r)$ can be bounded as follows:

$$\frac{1}{2}[C_{d,2\delta-1,2r} + C_{d,A,2r}] \leq C_d(\mathbb{D}_{n,\delta}, r) \leq C_{d,2\delta-1,2r}. \quad (8.3.7)$$

where $A = (2\delta + 1, 2\delta - 1, \dots, 2\delta - 1) \in \mathbb{R}^d$.

The proof of Lemma 8.3.2 is given in Appendix B, see Section 8.6.

In Figures 8.3 and 8.4, using the approximation given in (8.3.6) we study the tightness of the bounds given in (8.3.7). In these figures, the dashed red line, dashed blue line and solid black line depict the upper bound, the lower bound and the approximation for $C_d(\mathbb{D}_{n,\delta}, r)$ respectively. We see that the upper bound is very sharp across r and d ; this behaviour is not seen with the lower bound.

8.3.6 ‘Do not try to cover the vertices’

In this section, we theoretically support the authors recommendation ‘do not try to cover the vertices’ which was the main message of Chapter 7 but was based on numerical evidence. In other words, we will show on the example of the design $\mathbb{D}_{n,1/2}$ that in large dimensions the attempt to cover the whole cube rather than 0.999 of it leads to a dramatic increase of the required radius of the balls.

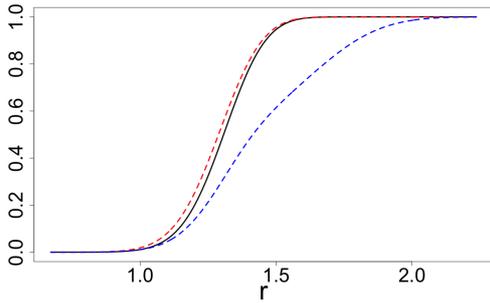


Figure 8.3: $C_d(\mathbb{D}_{n,\delta}, r)$ with upper and lower bounds: $d = 20, \delta = 1/2$

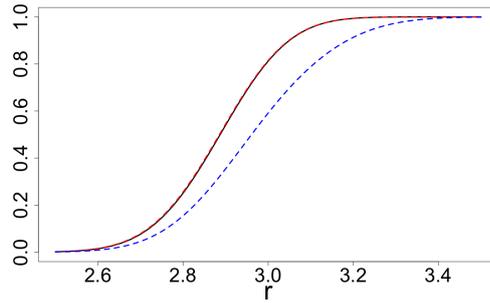


Figure 8.4: $C_d(\mathbb{D}_{n,\delta}, r)$ with upper and lower bounds: $d = 100, \delta = 1/2$

Theorem 8.3.2 *Let γ be fixed, $0 \leq \gamma \leq 1$. Consider $(1 - \gamma)$ -coverings of $[-1, 1]^d$ generated by the designs $\mathbb{D}_{n,\delta}$ and the associated normalized radii $R_{1-\gamma}(\mathbb{D}_{n,\delta})$, see (8.1.10). For any $0 < \gamma < 1$ and $0 \leq \delta \leq 1$, the limit of $R_{1-\gamma}(\mathbb{D}_{n,\delta})$, as $d \rightarrow \infty$, exists and achieves minimal value for $\delta = 1/2$. Moreover, $R_{1-\gamma}(\mathbb{D}_{n,1/2})/R_1(\mathbb{D}_{n,1/2}) \rightarrow 1/\sqrt{3}$ as $d \rightarrow \infty$, for any $0 < \gamma < 1$.*

Proof is given in Appendix C, see Section 8.7.

In Figures 8.5-8.6 using a solid red line we depict the approximation of $C_d(\mathbb{D}_{n,\delta}, r)$ as a function of R with $\delta = 1/2$, where we recall $r = 2\sqrt{d}n^{-1/d}R$. The vertical green line illustrates the value of $R_{0.999}$ and the vertical blue line depicts $R_1 = n^{1/d}r_1/(2\sqrt{d}) = n^{1/d}\sqrt{d+8}/(4\sqrt{d})$, where for design $\mathbb{D}_{n,1/2}$ one can easily calculate (using Lemma 8.3.1) that $r_1 = \sqrt{d+8}/2$. This corresponds to computing the distance between the point $\mathbf{1}/2 = (1/2, \dots, 1/2)$ and the furthest point in the Voronoi cell around $\mathbf{1}/2$ which is (using Theorem 8.2.1) the point $(-1, 1, \dots, 1)$. These figures clearly illustrate that as d increases, for all γ we have $R_{1-\gamma}/R_1$ slowly tending to $1/\sqrt{3}$ (that is, the ratio of the radius depicted with the green line and the radius shown with the blue line tends to $1/\sqrt{3}$).

In Figures 8.7-8.10, instead of plotting with respect to the normalised radius R , we have returned to the non-normalised radius r . In these figures the solid red line depicts the approximation of $C_d(\mathbb{D}_{n,\delta}, r)$ developed in Section 8.3.4 as a function of r with $\delta = 1/2$. The vertical green line indicates the value of $r_{0.999}$ and the vertical blue line depicts $r_1 = \sqrt{d+8}/2$. One can clearly see in these figures how much smaller $r_{0.999}$ is relative to r_1 , the radius guaranteeing the full coverage.

8.4 Numerical studies

For comparative purposes, we introduce another design which is one of the most popular designs (both, for quantization and covering) considered in applications:

Design \mathbb{S}_n : Z_1, \dots, Z_n are taken from a low-discrepancy Sobol's sequence on the cube $[-1, 1]^d$. (See Scheme 7 in Chapter 7 and references therein.)

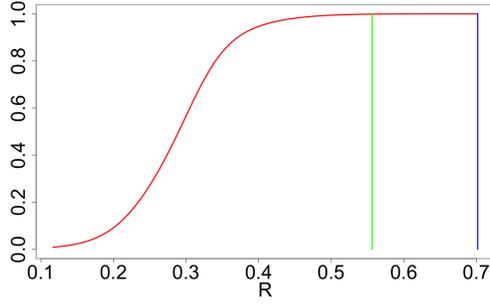


Figure 8.5: $C_d(\mathbb{D}_{n,\delta}, r)$ with $R_{0.999}$ and R_1 : $d = 5$

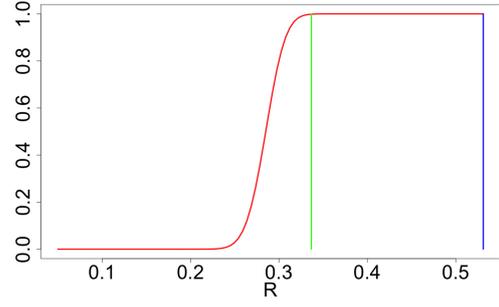


Figure 8.6: $C_d(\mathbb{D}_{n,\delta}, r)$ with $R_{0.999}$ and R_1 : $d = 50$

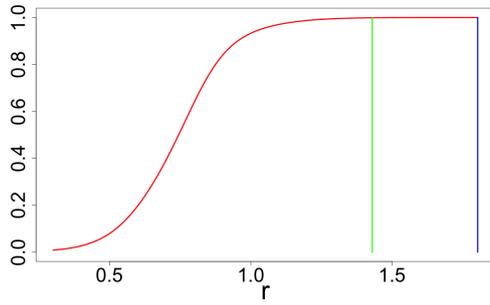


Figure 8.7: $C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 5$

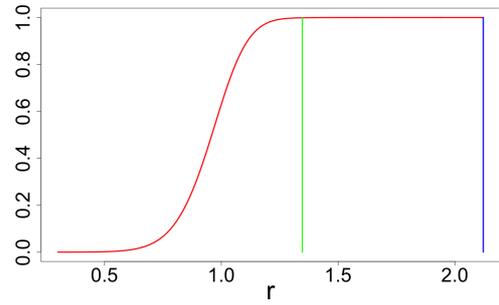


Figure 8.8: $C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 10$

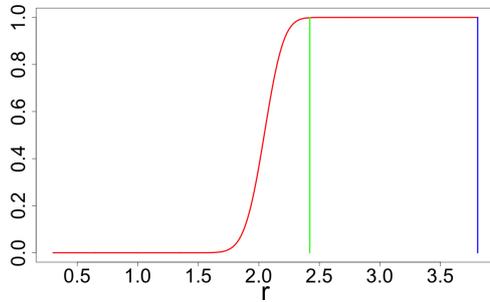


Figure 8.9: $C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 50$

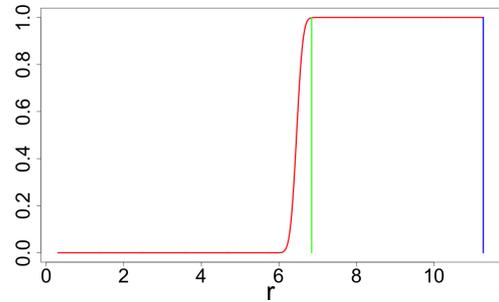


Figure 8.10: $C_d(\mathbb{D}_{n,\delta}, r)$ with $r_{0.999}$ and r_1 : $d = 500$

For constructing the design \mathbb{S}_n , we use the R-implementation provided in the ‘SobolSequence’ package, see [57]. For \mathbb{S}_n , we have set $n = 1024$ and $F2 = 10$ (an input parameter for the Sobol sequence function).

8.4.1 Quantization and weak covering comparisons

In Table 8.1, we compare the normalised mean squared quantization error $Q_d(\mathbb{Z}_n)$ defined in (8.1.6) across three designs: \mathbb{D}_{n,δ^*} with δ^* given in (8.2.11), $\mathbb{D}_n^{(0)}$ and \mathbb{S}_n . Table 8.1 and similar comparisons with the other designs have inspired us to formulate the conjecture in Section 8.1.7.

In Table 8.2, we compare the normalised statistic $R_{1-\gamma}$ introduced in (8.1.9),

	$d = 5$	$d = 7$	$d = 10$	$d = 15$	$d = 20$
$Q_d(\mathbb{D}_{n,\delta^*})$	0.0876	0.0827	0.0804	0.0798	0.0800
$Q_d(\mathbb{D}_n^{(0)})$	0.0833	0.0833	0.0833	0.0833	0.0833
$Q_d(\mathbb{S}_n)$	0.0988	0.1003	0.1022	0.1060	0.1086

Table 8.1: Normalised mean squared quantization error Q_d for four designs and different d .

where we have fixed $\gamma = 0.01$. For designs $\mathbb{D}_{n,\delta}$, $\mathbb{D}_{n,0.5}$ and $\mathbb{D}_n^{(0)}$ we have also included R_1 , the smallest normalised radius that ensures the full coverage. For design $\mathbb{D}_{n,\delta}$ and for each d , the optimal values of δ for this γ are provided in brackets and have been obtained using the approximation of Section 8.3.4.

	$d = 5$	$d = 7$	$d = 10$	$d = 15$	$d = 20$
$R_{1-\gamma}(\mathbb{D}_{n,\delta})$	0.4750 (0.54)	0.3992 (0.53)	0.3635 (0.52)	0.3483 (0.51)	0.3417 (0.50)
$R_{1-\gamma}(\mathbb{D}_{n,0.5})$	0.4765	0.4039	0.3649	0.3484	0.3417
$R_{1-\gamma}(\mathbb{D}_n^{(0)})$	0.4092	0.3923	0.3766	0.3612	0.3522
$R_{1-\gamma}(\mathbb{S}_n)$	0.4714	0.4528	0.4256	0.4074	0.3967

$R_1(\mathbb{D}_{n,\delta})$	0.6984 (0.54)	0.6555 (0.53)	0.6178 (0.52)	0.5856 (0.51)	0.5714 (0.50)
$R_1(\mathbb{D}_{n,0.5})$	0.7019	0.6629	0.6259	0.5912	0.5714
$R_1(\mathbb{D}_n^{(0)})$	0.5000	0.5000	0.5000	0.5000	0.5000

Table 8.2: Normalised statistic $R_{1-\gamma}$ across d with $\gamma = 0.01$ (value in brackets corresponds to optimal δ)

Let us make some remarks on Tables 8.1 and 8.2:

- In conjunction with Figure 8.1, Table 8.1 shows that for $d \geq 7$, the quantization for design \mathbb{D}_{n,δ^*} is superior over all other designs considered.
- For the weak covering statistic $R_{1-\gamma}$, the superiority of $\mathbb{D}_{n,\delta}$ over all other designs considered is seen for $d \geq 10$.
- For the design $\mathbb{D}_{n,\delta}$, the optimal value of δ minimizing $R_{1-\gamma}$ is not the same as the optimal δ for quantization.
- From one of the five remarks given in Section 8.2.3, the minimal value of $Q_d(\mathbb{D}_{n,\delta^*})$ is attained with $d = 15$. For $d > 15$, the quantity $Q_d(\mathbb{D}_{n,\delta^*})$ increases with d , slowly converging to $Q_d(\mathbb{D}_n^{(0)}) = 1/12$. This non-monotonic behaviour can be seen by looking at $d = 20$ in Table 8.1.
- Similar non-monotonic behaviour is not seen for the quantity $R_{1-\gamma}$, as $R_{1-\gamma}(\mathbb{D}_{n,\delta})$ monotonically decreases as d increases. Also, Theorem 8.3.2 and its proof imply the asymptotically optimal value of δ is $1/2$ and that $R_{1-\gamma}(\mathbb{D}_{n,\delta}) \rightarrow 1/(2\sqrt{3}) \cong 0.289$ as $d \rightarrow \infty$. The latter statement follows from the proof of Theorem 8.3.2 where it is shown that $r_{1-\gamma} = \frac{\sqrt{d}}{2\sqrt{3}}$. After normalisation, the result can be obtained.

8.4.2 Accuracy of covering approximation and dependence on δ

In this section, we assess the accuracy of the approximation of $C_d(\mathbb{D}_{n,\delta}, r)$ developed in Section 8.3.4 and the behaviour of $C_d(\mathbb{D}_{n,\delta}, r)$ as a function of δ . In Figures 8.11 – 8.16, the thick dashed black lines depict $C_d(\mathbb{D}_{n,\delta}, r)$ for several different choices of r ; these values are obtained via Monte Carlo simulations with 50,000 repetitions. The thinner solid lines depicts its approximation of Section 8.3.4. These figures show that the approximation is extremely accurate for all r , δ and d ; we emphasise that the approximation remains accurate even for very small dimensions like $d = 3$. These figures also demonstrate the δ -effect saying that a significantly more efficient weak covering can be achieved with a suitable choice of δ . This is particularly evident in higher dimensions, see Figures 8.15 and 8.16.

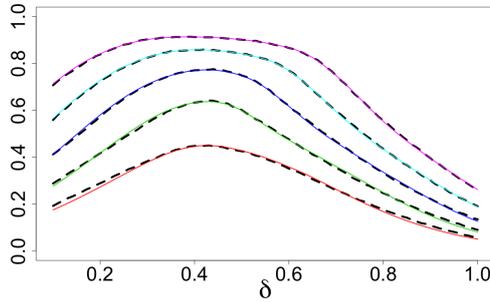


Figure 8.11: $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 3, r$ from 0.6 to 1 increasing by 0.1

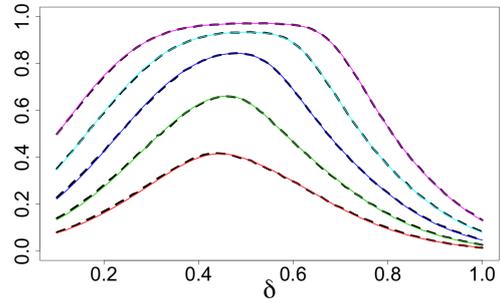


Figure 8.12: $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 5, r$ from 0.7 to 1.1 increasing by 0.1

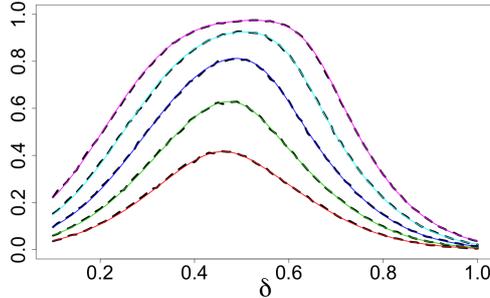


Figure 8.13: $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 7, r$ from 0.8 to 1.1 increasing by 0.075

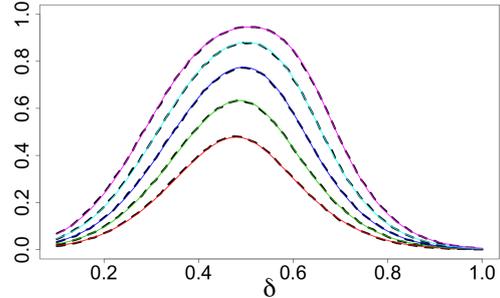


Figure 8.14: $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 10, r$ from 0.95 to 1.15 increasing by 0.05

Figures 8.17 and 8.18 illustrate Theorem 8.3.2 and show the rate of convergence of the covering radii as d increases. Let the probability density function $f(r)$ be defined by $dC_d(\mathbb{D}_{n,\delta}, r) = f(r)dr$, where $C_d(\mathbb{D}_{n,\delta}, r)$ as a function of r is viewed as the c.d.f. of the r.v. $r = \varrho(X, \mathbb{Z}_n)$, see Section 8.1.3. Trivial calculations yield that the density for the normalised form of $\varrho(X, \mathbb{Z}_n)$ using (8.1.7) is $p_d(R) := 2\sqrt{d}n^{-1/d}f\left(2\sqrt{d}n^{-1/d}R\right)$. In Figure 8.17, we depict the density $p_d(\cdot)$ for $d = 5, 10$ and 20 with blue, red and black lines respectively. The respective c.d.f.'s $\int_0^R p_d(\tau)d\tau$ are shown in Figure 8.18 under the same colour scheme.

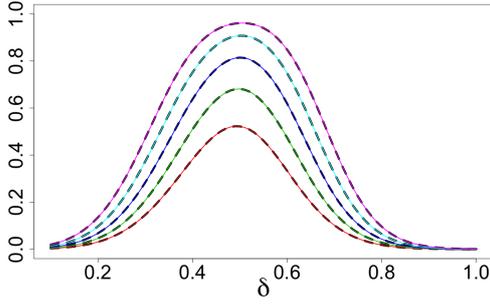


Figure 8.15: $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 15$, r from 1.15 to 1.35 increasing by 0.05

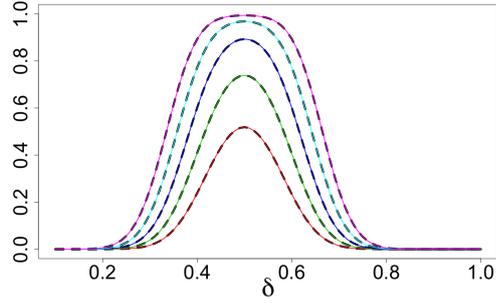


Figure 8.16: $C_d(\mathbb{D}_{n,\delta}, r)$ and its approximation: $d = 50$, r from 2.05 to 2.35 increasing by 0.075

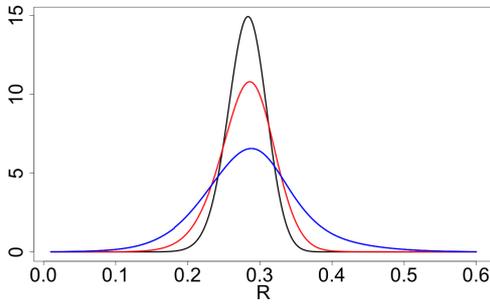


Figure 8.17: Densities $f_d(R)$ for the design \mathbb{D}_{n,δ^*} ; $d = 5, 10, 20$

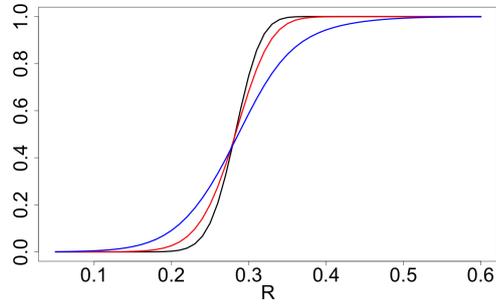


Figure 8.18: c.d.f.'s of R for the design \mathbb{D}_{n,δ^*} ; $d = 5, 10, 20$

8.4.3 Stochastic dominance

In Figures 8.19 and 8.20, we depict the c.d.f.'s for the normalized distance $n^{1/d}\varrho(X, \mathbb{Z}_n)/(2\sqrt{d})$ for two designs: \mathbb{D}_{n,δ^*} in red, and $\mathbb{D}_n^{(0)}$ in black. We can see that the design \mathbb{D}_{n,δ^*} stochastically dominates the design $\mathbb{D}_n^{(0)}$ for $d = 10$ but for $d = 5$ there seem to be a reverse domination; this is in line with findings from Sections 8.2.3 and 8.4.1, see e.g. Figure 8.1, Tables 8.1 and 8.2.

In Figure 8.21, we depict the c.d.f.'s for the normalized distance $n^{1/d}\varrho(X, \mathbb{Z}_n)/(2\sqrt{d})$ for design $\mathbb{D}_n^{(0)}$ (in red) and design \mathbb{S}_n (in black). We can see that for $d = 5$, the design $\mathbb{D}_n^{(0)}$ stochastically dominates the design \mathbb{S}_n . The style of Figure 8.22 is the same as figure Figure 8.21, however we set $d = 10$ and the design $\mathbb{D}_n^{(0)}$ is replaced with the design \mathbb{D}_{n,δ^*} . Here we see a very clear stochastic dominance of the design \mathbb{D}_{n,δ^*} over the design \mathbb{S}_n . All findings are consistent with findings from Section 8.4.1, see Tables 8.1 and 8.2.

8.5 Appendix A: Proof of Theorem 8.3.1

In view of (8.3.1), $C_d(\mathbb{D}_{n,\delta}, r) = V_{d,\delta,r}$ for all $0 \leq \delta \leq 1$ and $r \geq 0$ and we shall derive expressions for $V_{d,\delta,r}$ rather than $C_d(\mathbb{D}_{n,\delta}, r)$.

Case(a) : $r \leq \delta$.

Let $Y = (y_1, y_2, \dots, y_d)$, where y_i ($i = 1, 2, \dots, d$) are i.i.d. random variables

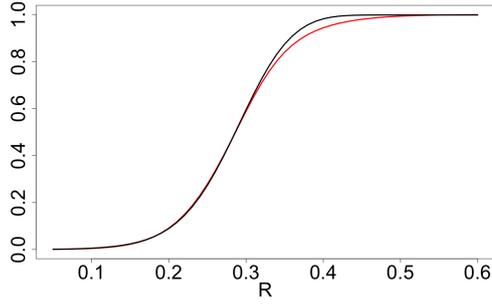


Figure 8.19: $d = 5$: design $\mathbb{D}_n^{(0)}$ seems to stochastically dominate \mathbb{D}_{n,δ^*}

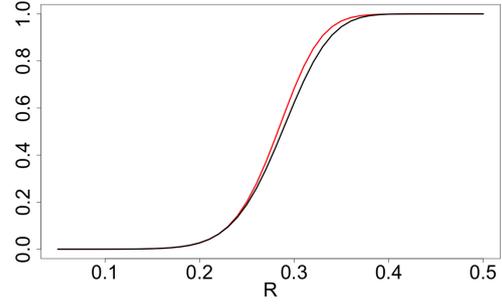


Figure 8.20: $d = 10$: design \mathbb{D}_{n,δ^*} stochastically dominates design $\mathbb{D}_n^{(0)}$

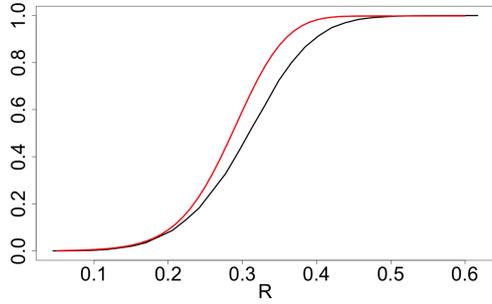


Figure 8.21: $d = 5$: design $\mathbb{D}_n^{(0)}$ stochastically dominates design \mathbb{S}_n

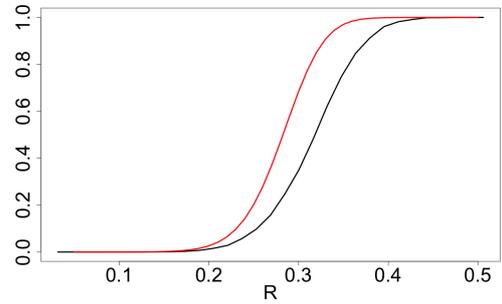


Figure 8.22: $d = 10$: design \mathbb{D}_{n,δ^*} stochastically dominates design \mathbb{S}_n

with uniform distribution on $[0, 1]$. Then for Case (a) we have:

$$\begin{aligned} V_{d,\delta,r} &= \frac{1}{2} \text{vol} \left(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap \left\{ X \in \mathbb{R}^d : 0 \leq x_i \leq 1, i = 1, 2, \dots, d \right\} \right) \\ &= \frac{1}{2} \Pr \{ \|Y - \boldsymbol{\delta}\| \leq r \} \\ &= \frac{1}{2} \Pr \left\{ \sum_{i=1}^d (y_i - \delta)^2 \leq r^2 \right\}. \end{aligned}$$

By making the substitution $y'_i = y_i - \frac{1}{2}$, we have y'_i are i.i.d. random variables with uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$, $i = 1, 2, \dots, d$. Therefore

$$V_{d,\delta,r} = \frac{1}{2} \Pr \left\{ \sum_{i=1}^d \left(y'_i + \frac{1}{2} - \delta \right)^2 \leq r^2 \right\} = \frac{1}{2} C_{d,\delta-\frac{1}{2},r}^{(1/2)}.$$

Whence, by using relation (8.3.2) with $\beta = 1/2$, we obtain

$$V_{d,\delta,r} = \frac{1}{2} C_{d,2\delta-1,2r}.$$

Case(b): $\delta \leq r \leq 1 + \delta$

Using (8.2.4) we obtain

$$V_{d,\delta,r} = \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + d \cdot \text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U_1) \right].$$

The first quantity in the brackets has been considered in case (a) and it is simply $C_{d,2\delta-1,2r}$. Therefore we aim to reformulate the second quantity within the brackets, $\mathcal{V} := \text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap U_1)$, in a probabilistic setting. Let $Y = (y_1, y_2, \dots, y_d)$, where y_1 has a uniform distribution on $[-1, 0]$ and y_i are i.i.d. random variables with a uniform distribution on $[|y_1|, 1]$, $i = 2, \dots, d$. By conditioning on y_1 and invoking the law of total probability, we obtain

$$\begin{aligned} \mathcal{V} &= \\ &= \int_{-1}^0 \Pr \{ \|Y - \boldsymbol{\delta}\| \leq r \mid y_1 \} \cdot \text{vol}(\{|y_1| \leq y_i \leq 1, i \neq 1\}) dy_1 \\ &= \int_{\delta-r}^0 \Pr \{ \|Y - \boldsymbol{\delta}\| \leq r \mid y_1 \} \cdot (1 - |y_1|)^{d-1} dy_1. \end{aligned} \quad (8.5.1)$$

The limits of integration have changed from $[-1, 0]$ to $[\delta-r, 0]$ since for $\delta \leq r \leq 1+\delta$ we have

$$\Pr \{ \|Y - \boldsymbol{\delta}\| \leq r \mid y_1 \} = 0 \text{ for } y_1 < \delta - r.$$

By focusing on the integrand in (8.5.1), we have:

$$\begin{aligned} \mathcal{P} &:= \Pr \{ \|Y - \boldsymbol{\delta}\| \leq r \mid y_1 \} \\ &= \Pr \left\{ \sum_{i=1}^d (y_i - \delta)^2 \leq r^2 \mid y_1 \right\} \\ &= \Pr \left\{ \sum_{i=2}^d (y_i - \delta)^2 \leq r^2 - (y_1 - \delta)^2 \mid y_1 \right\}. \end{aligned}$$

For $i = 2, 3, \dots, d$, by making the substitution $y'_i = y_i - \frac{1+|y_1|}{2}$, we have y'_i are i.i.d. with uniform distribution on $[\frac{1}{2}(|y_1| - 1), \frac{1}{2}(1 - |y_1|)]$. Let $\delta' = (1 - |y_1|)/2$. This results in:

$$\begin{aligned} \mathcal{P} &= \Pr \left\{ \sum_{i=2}^d \left(y'_i - \delta + \frac{1+|y_1|}{2} \right)^2 \leq r^2 - (y_1 - \delta)^2 \mid y_1 \right\} \\ &= C_{d-1, \frac{2\delta-1-|y_1|}{2}, \sqrt{r^2-(y_1-\delta)^2}}^{(\delta')}. \end{aligned}$$

Using relation (8.3.2) with $\beta = \delta' = (1 - |y_1|)/2$ we obtain:

$$\mathcal{P} = C_{d-1, \frac{2\delta-1-|y_1|}{1-|y_1|}, \frac{2\sqrt{r^2-(y_1-\delta)^2}}{1-|y_1|}},$$

whence

$$\mathcal{V} = \int_{\delta-r}^0 C_{d-1, \frac{2\delta-1-|y_1|}{1-|y_1|}, \frac{2\sqrt{r^2-(y_1-\delta)^2}}{1-|y_1|}} \cdot (1 - |y_1|)^{d-1} dy_1$$

and we conclude:

$$\begin{aligned} V_{d,\delta,r} &= \frac{1}{2} \left[C_{d,2\delta-1,2r} + d \int_{\delta-r}^0 C_{d-1, \frac{2\delta-1-|x|}{1-|x|}, \frac{2\sqrt{r^2-(x-\delta)^2}}{1-|x|}} (1-|x|)^{d-1} dx \right] \quad (8.5.2) \\ &= \frac{1}{2} \left[C_{d,2\delta-1,2r} + d \int_0^{r-\delta} C_{d-1, \frac{2\delta-1-x}{1-x}, \frac{2\sqrt{r^2-(x+\delta)^2}}{1-x}} (1-x)^{d-1} dx \right]. \end{aligned}$$

Case(c): $r \geq 1 + \delta$:

Case (c) is almost identical to Case (b), with the only change occurring within the lower limit of integration in (8.5.2). Since y_1 is constrained within $[-1, 0]$ and $r \geq 1 + \delta$, the lower limit of the integral remains at -1 for all $r \geq 1 + \delta$. Since the steps are almost identical to Case (b), they are omitted and we simply conclude:

$$V_{d,\delta,r} = \frac{1}{2} \left[C_{d,2\delta-1,2r} + d \int_0^1 C_{d-1, \frac{2\delta-1-x}{1-x}, \frac{2\sqrt{r^2-(x+\delta)^2}}{1-x}} (1-x)^{d-1} dx \right].$$

□

8.6 Appendix B: Proof of Lemma 8.3.2

(a) Let us first prove the upper bound in (8.3.7). Consider the set U_j defined in (8.2.5) and the associated set

$$U'_j = \left\{ X = (x_1, x_2, \dots, x_d) \in [0, 1]^d : |x_j| \leq x_k \leq 1 \text{ for all } k \neq j \right\} \subset C_0.$$

We have $\text{vol}(U_j) = \text{vol}(U'_j) = 1/d$ and

$$V(\delta) = C_0 \cup \left[\bigcup_{j=1}^d U_j \right], \quad \bigcup_{j=1}^d U'_j = C_0 = [0, 1]^d \quad (8.6.1)$$

Let us prove that for any $r \geq 0$ we have $\text{vol}(U_j \cap \mathcal{B}_d(\delta, r)) \leq \text{vol}(U'_j \cap \mathcal{B}_d(\delta, r))$. With any point $X = (x_1, x_2, \dots, x_d) \in U'_j$, we associate the point $X' = (-x_1, x_2, \dots, x_d) \in U_1$ by simply changing the sign in the first component. For these two points, we have

$$\|\delta - X\|^2 = (x_1 - \delta)^2 + \sum_{k=2}^d (x_k - \delta)^2 < (-x_1 - \delta)^2 + \sum_{k=2}^d (x_k - \delta)^2 = \|\delta - X'\|^2$$

Therefore, $\|\delta - X\|^2 \leq r \Rightarrow \|\delta - X'\|^2 \leq r$ yielding:

$$\text{vol}(U_j \cap \mathcal{B}_d(\delta, r)) \leq \text{vol}(U'_j \cap \mathcal{B}_d(\delta, r)). \quad (8.6.2)$$

To prove the upper bound in (8.3.7) for all r we must consider two cases: $r \leq \delta$ and $r \geq \delta$.

For $r \leq \delta$, we clearly have

$$V_{d,\delta,r} = \frac{1}{2} C_{d,2\delta-1,2r} \leq C_{d,2\delta-1,2r}$$

For $r \geq \delta$, using (8.6.2) we have

$$\begin{aligned}
V_{d,\delta,r} &= \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\delta, r) \cap C_0) + d \cdot \text{vol}(\mathcal{B}_d(\delta, r) \cap U_1) \right] \\
&\leq \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\delta, r) \cap C_0) + d \cdot \text{vol}(\mathcal{B}_d(\delta, r) \cap U'_1) \right] \\
&= \text{vol}(\mathcal{B}_d(\delta, r) \cap C_0) \\
&= C_{d,2\delta-1,2r}
\end{aligned}$$

and hence the upper bound in (8.3.7).

(b) Consider now the lower bound in (8.3.7). With the set U_j we now associate the set

$$\begin{aligned}
V_j &= \\
\left\{ \tilde{X} = (x_1, \dots, x_d) : -1 \leq x_1 \leq 0, 0 \leq x_m \leq 1 \text{ (for } m > 1), |x_j| \leq |x_k| \leq 1 \text{ for } k \neq j \right\}.
\end{aligned}$$

With any point $X = (x_1, x_2, \dots, x_d) \in U_j$ (here x_j is negative and $|x_j| \leq |x_k| \leq 1$ for $k \neq j$) we associate point $\tilde{X} = (-x_1, x_2, \dots, x_{j-1}, -x_j, x_{j+1}, \dots, x_d) \in V_j$ by changing sign in the first and j -th component of $X \in U_j$.

Setting without loss of generality $j = 2$, we have for these two points:

$$\begin{aligned}
\|\delta - X\|^2 &= (x_1 - \delta)^2 + (x_2 - \delta)^2 + \sum_{k=3}^d (x_k - \delta)^2 \\
&\leq (-x_1 - \delta)^2 + (-x_2 - \delta)^2 + \sum_{k=3}^d (x_k - \delta)^2 = \|\delta - \tilde{X}\|^2,
\end{aligned}$$

where the inequality follows from the inequalities $x_1 \geq 0$, $x_2 < 0$ and $|x_2| < x_1$ containing in the definition of U_2 .

Therefore, $\|\delta - X\|^2 \leq r \Rightarrow \|\delta - \tilde{X}\|^2 \leq r$ implying:

$$\text{vol}(U_j \cap \mathcal{B}_d(\delta, r)) \geq \text{vol}(V_j \cap \mathcal{B}_d(\delta, r)). \quad (8.6.3)$$

To prove the lower bound in (8.3.7) for all r we must consider two cases: $r \leq \delta$ and $r \geq \delta$.

For $r \geq \delta$, using (8.6.3) and the neighbouring cube C_1 defined in the proof of Theorem 8.2.1, we have

$$\begin{aligned}
V_{d,\delta,r} &= \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\delta, r) \cap C_0) + \sum_{i=1}^d \text{vol}(\mathcal{B}_d(\delta, r) \cap U_i) \right] \\
&\geq \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\delta, r) \cap C_0) + \text{vol}(\mathcal{B}_d(\delta, r) \cap U_1) + \sum_{i=2}^d (\mathcal{B}_d(\delta, r) \cap V_i) \right] \\
&= \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\delta, r) \cap C_0) + \text{vol}(\mathcal{B}_d(\delta, r) \cap C_1) \right].
\end{aligned}$$

Let y_1 have uniform distribution on $[-1, 0]$ and y_i have uniform distribution on $[0, 1]$ for $i \geq 2$. We have

$$\text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1) = \Pr \{ \|Y - \boldsymbol{\delta}\| \leq r \} = \Pr \left\{ \sum_{i=1}^d (y_i - \delta)^2 \leq r^2 \right\}.$$

For y_1 , make the substitution $y'_1 = y_1 + \frac{1}{2}$, we have y'_1 has uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$. For $i = 2, 3, \dots, d$, by making the substitution $y'_i = y_i - \frac{1}{2}$, we have y'_i are i.i.d. with uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$. This results in:

$$\text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1) = \Pr \left\{ \left(y'_1 - \frac{1}{2} - \delta \right)^2 + \sum_{i=2}^d (y'_i + \frac{1}{2} - \delta)^2 \leq r^2 \right\} = C_{d, B, r}^{(1/2)},$$

where $B = (\delta + \frac{1}{2}, \delta - \frac{1}{2}, \dots, \delta - \frac{1}{2})$. By then using relation (8.3.2) we obtain:

$$\text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1) = C_{d, 2B, 2r}$$

and hence we can conclude:

$$\begin{aligned} V_{d, \boldsymbol{\delta}, r} &\geq \frac{1}{2} \left[\text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_0) + \text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1) \right] \\ &= \frac{1}{2} [C_{d, 2\boldsymbol{\delta} - 1, 2r} + C_{d, A, 2r}], \end{aligned}$$

where $A = 2B = (2\delta + 1, 2\delta - 1, \dots, 2\delta - 1)$.

For $r \leq \delta$, since $\text{vol}(\mathcal{B}_d(\boldsymbol{\delta}, r) \cap C_1) = C_{d, A, 2r} = 0$, we have

$$V_{d, \boldsymbol{\delta}, r} = \frac{1}{2} [C_{d, 2\boldsymbol{\delta} - 1, 2r} + C_{d, A, 2r}]$$

and hence the lower bound in (8.3.7). \square

8.7 Appendix C: Proof of Theorem 8.3.2.

Before proving Theorem 8.3.2, we prove three auxiliary lemmas.

Lemma 8.7.1 *Let $r = r_{\alpha, d} = \alpha\sqrt{d}$ with $\alpha \geq 0$ and $Z_{a, b; d} = (a, b, b, \dots, b) \in \mathbb{R}^d$. Then the limit $\lim_{d \rightarrow \infty} C_{d, Z_{a, b; d}, 2r}$ exists and*

$$\lim_{d \rightarrow \infty} C_{d, Z_{a, b; d}, 2r} = \begin{cases} 0 & \text{if } \alpha < \frac{1}{2}\sqrt{\frac{1}{3} + b^2} \\ 1/2 & \text{if } \alpha = \frac{1}{2}\sqrt{\frac{1}{3} + b^2} \\ 1 & \text{if } \alpha > \frac{1}{2}\sqrt{\frac{1}{3} + b^2} \end{cases}.$$

Proof. Define

$$t_\alpha = \frac{\sqrt{3}(d(4\alpha^2 - b^2 - 1/3) + b^2 - a^2)}{2\sqrt{a^2 + (d-1)b^2 + d/15}}.$$

Let u have the uniform distribution on $[-1, 1]$ and $z \in \mathbb{R}$. As the r.v. $\eta_z := (u - z)^2$ is concentrated on a finite interval, for finite a and b the quantities of $\rho_a := \mathbb{E}(|\eta_a - a^2 - \frac{1}{3}|^3)$ and $\rho_b := \mathbb{E}(|\eta_b - b^2 - \frac{1}{3}|^3)$ are bounded. By applying Berry-Esseen theorem, see [9], to $C_{d, Z_{a,b}, 2r}$, there exists some constant C such that

$$-\frac{C \cdot \max\{\rho_a/\sigma_a^2, \rho_b/\sigma_b^2\}}{(\sigma_a^2 + (d-1)\sigma_b^2)^{1/2}} + \Phi(t_\alpha) \leq C_{d, Z_{a,b}, 2r} \leq \Phi(t_\alpha) + \frac{C \cdot \max\{\rho_a/\sigma_a^2, \rho_b/\sigma_b^2\}}{(\sigma_a^2 + (d-1)\sigma_b^2)^{1/2}},$$

where $\sigma_a^2 = \text{var}(\eta_a)$ and $\sigma_b^2 = \text{var}(\eta_b)$. By the squeeze theorem, it is clear that if $4\alpha^2 - b^2 - 1/3 > 0$ and hence $\alpha > \frac{1}{2}\sqrt{\frac{1}{3} + b^2}$, then $C_{d, Z_{a,b}, 2r} \rightarrow 1$ as $d \rightarrow \infty$. If $\alpha < \frac{1}{2}\sqrt{\frac{1}{3} + b^2}$, then $C_{d, Z_{a,b}, 2r} \rightarrow 0$ as $d \rightarrow \infty$. If $\alpha = \frac{1}{2}\sqrt{\frac{1}{3} + b^2}$, then $C_{d, Z_{a,b}, 2r} \rightarrow 1/2$ as $d \rightarrow \infty$. \square

Lemma 8.7.2 *Let $r = \alpha\sqrt{d}$. Then for $\boldsymbol{\delta} = (\delta, \delta, \dots, \delta)$, we have:*

$$\lim_{d \rightarrow \infty} V_{d, \boldsymbol{\delta}, r} = \lim_{d \rightarrow \infty} C_{d, 2\boldsymbol{\delta} - 1, 2r} = \begin{cases} 0 & \text{if } \alpha < \frac{\sqrt{1/3 + (2\delta - 1)^2}}{2} \\ 1/2 & \text{if } \alpha = \frac{\sqrt{1/3 + (2\delta - 1)^2}}{2} \\ 1 & \text{if } \alpha > \frac{\sqrt{1/3 + (2\delta - 1)^2}}{2} \end{cases}$$

Proof. Using Lemma 8.7.1 with $Z_{a,b} = A = (2\delta + 1, 2\delta - 1, \dots, 2\delta - 1)$, we obtain:

$$\lim_{d \rightarrow \infty} C_{d, A, 2r} = \lim_{d \rightarrow \infty} C_{d, 2\boldsymbol{\delta} - 1, 2r} = \begin{cases} 0 & \text{if } \alpha < \frac{\sqrt{1/3 + (2\delta - 1)^2}}{2} \\ 1/2 & \text{if } \alpha = \frac{\sqrt{1/3 + (2\delta - 1)^2}}{2} \\ 1 & \text{if } \alpha > \frac{\sqrt{1/3 + (2\delta - 1)^2}}{2} \end{cases}$$

By then applying the squeeze theorem to the bounds in Lemma 8.3.2 using the fact from Lemma 8.3.1 we have $V_{d, \boldsymbol{\delta}, r} = C_d(\mathbb{Z}_n, r)$, we obtain the result. \square

To determine the value of r that leads to the full coverage, we utilise the following simple lemma.

Lemma 8.7.3 *For design $\mathbb{D}_{n, \boldsymbol{\delta}}$, the smallest value of r that ensures a complete covering of $[-1, 1]^d$ as $d \rightarrow \infty$ is:*

$$r_1 = \begin{cases} (1 - \delta)\sqrt{d} & \text{if } \delta \leq 1/2 \\ \delta\sqrt{d} & \text{if } \delta > 1/2 \end{cases}.$$

Proof. Using Lemma 8.3.1, we can, without loss of generality, focus on the complete covering of $V(\boldsymbol{\delta})$, which is given in (8.6.1). In this Voronoi cell, the only possible candidate points that could lead to r_1 are $(-1, 1, \dots, 1)$, $(0, 0, \dots, 0)$ and $(1, 1, \dots, 1)$. For these three points, for a fixed d we have:

$$\begin{aligned} \|\boldsymbol{\delta} - (-1, 1, \dots, 1)\| &= \sqrt{(-1 - \delta)^2 + (d-1)(1 - \delta)^2} \\ \|\boldsymbol{\delta} - (0, 0, \dots, 0)\| &= \delta\sqrt{d} \\ \|\boldsymbol{\delta} - (1, 1, \dots, 1)\| &= \sqrt{d(1 - \delta)^2}. \end{aligned}$$

Therefore, as $d \rightarrow \infty$, depending on the value of δ we either have $r_1 = \delta\sqrt{d}$ or $r_1 = (1 - \delta)\sqrt{d}$. We have $\delta\sqrt{d} \leq (1 - \delta)\sqrt{d}$ when $\delta \leq 1/2$ and hence for $\delta \leq 1/2$, we have $r_1 = (1 - \delta)\sqrt{d}$. Likewise, for $\delta > 1/2$ we obtain $r_1 = (1 - \delta)\sqrt{d}$. \square

Proof of Theorem 8.3.2.

From Lemma 8.7.2, it is clear that the smallest α and hence r is attained with $\delta = 1/2$. Moreover, Lemma 8.7.2 provides:

$$\lim_{d \rightarrow \infty} V_{d,1/2,r} = \lim_{d \rightarrow \infty} C_{d,0,2r} = \begin{cases} 0 & \text{if } \alpha < \frac{1}{2\sqrt{3}} \\ 1/2 & \text{if } \alpha = \frac{1}{2\sqrt{3}} \\ 1 & \text{if } \alpha > \frac{1}{2\sqrt{3}} \end{cases} \quad (8.7.1)$$

meaning for any $0 < \gamma < 1$, $r_{1-\gamma} = \frac{\sqrt{d}}{2\sqrt{3}}$. By then applying Lemma 8.7.3 with $\delta = 1/2$, we obtain $r_1 = \sqrt{d}/2$ and hence $r_{1-\gamma}/r_1 \rightarrow 1/\sqrt{3}$ as $d \rightarrow \infty$. The quantity on the r.h.s. of (8.7.1) is a pointwise limit and is not a distribution function (it is not right continuous). If we consider convergence in distributions, we converge to the distribution function

$$F(\alpha) := \begin{cases} 0 & \text{if } \alpha < \frac{1}{2\sqrt{3}} \\ 1 & \text{if } \alpha \geq \frac{1}{2\sqrt{3}} \end{cases}$$

since for all α at which $F(\alpha)$ is continuous, we have $\lim_{d \rightarrow \infty} V_{d,1/2,r} = F(\alpha)$. \square

Chapter 9

Non-lattice covering and quantization of high dimensional sets

Abstract

The main problem considered in this chapter is construction and theoretical study of efficient n -point coverings of a d -dimensional cube $[-1, 1]^d$. Targeted values of d are between 5 and 50; n can be in hundreds or thousands and the designs (collections of points) are nested. This chapter is a continuation of Chapter 7, where we have theoretically investigated several simple schemes and numerically studied many more. In this chapter, we extend the theoretical constructions of Chapter 7 for studying the designs which were found to be superior to the ones theoretically investigated in Chapter 7. In this chapter, we extend our constructions for new construction schemes which provide even better coverings (in the class of nested designs) than the ones numerically found in Chapter 7. In view of a close connection of the problem of quantization to the problem of covering, we extend our theoretical approximations and practical recommendations to the problem of construction of efficient quantization designs in a cube $[-1, 1]^d$. In the last section, we discuss the problems of covering and quantization in a d -dimensional simplex. The content of this chapter has been published in [86].

9.1 Introduction

The problem of the main importance in this chapter is the following problem of covering a cube $[-1, 1]^d$ by n balls. Let Z_1, \dots, Z_n be a collection of points in \mathbb{R}^d and $\mathcal{B}_d(Z_j, r) = \{Z : \|Z - Z_j\| \leq r\}$ be the Euclidean balls of radius r centered at Z_j ($j = 1, \dots, n$). The dimension d , the number of balls n and their radius r could be arbitrary.

We are interested in choosing the locations of the centers of the balls Z_1, \dots, Z_n so that the union of the balls $\cup_j \mathcal{B}_d(Z_j, r)$ covers the largest possible proportion of the cube $[-1, 1]^d$. More precisely, we are interested in choosing a collection of points

(called ‘design’) $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ so that

$$C_d(\mathbb{Z}_n, r) := \text{vol}([-1, 1]^d \cap \mathcal{B}_d(\mathbb{Z}_n, r)) / 2^d \quad (9.1.1)$$

is as large as possible (given n , r and the freedom we are able to use in choosing Z_1, \dots, Z_n). Here $\mathcal{B}_d(\mathbb{Z}_n, r)$ is the union of the balls

$$\mathcal{B}_d(\mathbb{Z}_n, r) = \bigcup_{j=1}^n \mathcal{B}_d(Z_j, r) \quad (9.1.2)$$

and $C_d(\mathbb{Z}_n, r)$ is the proportion of the cube $[-1, 1]^d$ covered by $\mathcal{B}_d(\mathbb{Z}_n, r)$. If $Z_j \in \mathbb{Z}_n$ are random then we shall consider $\mathbb{E}_{\mathbb{Z}_n} C_d(\mathbb{Z}_n, r)$, the expected value of the proportion (9.1.1); for simplicity of notation, we will drop $\mathbb{E}_{\mathbb{Z}_n}$ while referring to $\mathbb{E}_{\mathbb{Z}_n} C_d(\mathbb{Z}_n, r)$.

For a design \mathbb{Z}_n , recall from Chapter 7 that its covering radius is defined by $\text{CR}(\mathbb{Z}_n) = \max_{X \in \mathcal{C}_d} \min_{Z_j \in \mathbb{Z}_n} \|X - Z_j\|$. If $r = \text{CR}(\mathbb{Z}_n)$, then $C_d(\mathbb{Z}_n, r)$ defined in (9.1.1) is equal to 1, and the whole cube \mathcal{C}_d gets covered by the balls. However, like Chapter 7 we are only interested in reaching the values like 0.95 or 0.99, when only a large part of the ball is covered. This is because the computation of $\text{CR}(\mathbb{Z}_n)$ is too difficult for large d . We refer the reader to the Introduction of Chapter 7 for discussions about why determining an optimal covering is an important problem.

We will say that $\mathcal{B}_d(\mathbb{Z}_n, r)$ makes a $(1 - \gamma)$ -covering of $[-1, 1]^d$ if

$$C_d(\mathbb{Z}_n, r) = 1 - \gamma; \quad (9.1.3)$$

the corresponding value of r will be called $(1 - \gamma)$ -covering radius and denoted $r_{1-\gamma}$ or $r_{1-\gamma}(\mathbb{Z}_n)$. If $\gamma = 0$ then the $(1 - \gamma)$ -covering becomes the full covering and 1-covering radius $r_1(\mathbb{Z}_n)$ becomes the covering radius $\text{CR}(\mathbb{Z}_n)$. The problem of construction of efficient designs with smallest possible $(1 - \gamma)$ -covering radius (with some small $\gamma > 0$) will be referred to as the problem of weak covering.

Two strong arguments why the problem of weak covering could be even more practically important than the problem of full covering are as follows.

- Numerical checking of weak covering (with an approximate value of γ) is straightforward while numerical checking of the full covering is practically impossible, if d is large enough.
- For a given design \mathbb{Z}_n , $C_d(\mathbb{Z}_n, r)$ defined in (9.1.1) and considered as a function of r , is a cumulative distribution function (c.d.f.) of the random variable (r.v.) $\varrho(U, \mathbb{Z}_n) = \min_{Z_i \in \mathbb{Z}_n} \|U - Z_i\|$, where U is a random vector uniformly distributed on $[-1, 1]^d$, see (9.7.1) below. The covering radius $\text{CR}(\mathbb{Z}_n)$ is the upper bound of this r.v. while in view of (9.1.3), $r_{1-\gamma}(\mathbb{Z}_n)$ is the $(1 - \gamma)$ -quantile. Many practically important characteristics of designs such as quantization error considered in Section 9.7 are expressed in terms of the whole c.d.f. $C_d(\mathbb{Z}_n, r)$ and their dependence on the upper bound $\text{CR}(\mathbb{Z}_n)$ is marginal. As shown in Section 9.7.5, numerical studies indicate that comparison of designs on the base of their weak coverage properties is very similar to quantization error comparisons, but this may not be true for comparisons with respect to

$\text{CR}(\mathbb{Z}_n)$. This phenomenon is similar to the well-known fact in the theory of space covering by lattices (see surveys [133, 134] and the excellent book [19]), where best lattice coverings of space are often poor quantizers and vice-versa. Moreover, Figures 9.1–9.2 below show that $\text{CR}(\mathbb{Z}_n)$ may give a totally inadequate impression about the c.d.f. $C_d(\mathbb{Z}_n, r)$ and could be much larger than $r_{1-\gamma}(\mathbb{Z}_n)$ with very small $\gamma > 0$.

In Figures 9.1–9.2 we consider two simple designs for which we plot their c.d.f. $C_d(\cdot, r)$, black line, and also indicate the location of the $r_1 = \text{CR}$ and $r_{0.999}$ by vertical red and green line respectively. In Figure 9.1, we take $d = 10$, $n = 512$ and use a 2^{d-1} design of maximum resolution concentrated at the points¹ $(\pm 1/2, \dots, \pm 1/2) \in \mathbb{R}^d$ as design \mathbb{Z}_n ; this design is a particular case of Design 4 of Section 9.8 and can be defined for any $d > 2$. In Figure 9.2, we keep $d = 10$ but take the full factorial 2^d design with $m = 2^d$ points, again concentrated at the points $(\pm 1/2, \dots, \pm 1/2)$; denote this design \mathbb{Z}'_m .

For both designs, it is very easy to analytically compute their covering radii (for any $d > 2$): $\text{CR}(\mathbb{Z}_n) = \sqrt{d+8}/2$ and $\text{CR}(\mathbb{Z}'_m) = \sqrt{d}/2$; for $d = 10$ this gives $\text{CR}(\mathbb{Z}_n) \simeq 2.1213$ and $\text{CR}(\mathbb{Z}'_m) \simeq 1.58114$. The values of $r_{0.999}$ are: $r_{0.999}(\mathbb{Z}_n) \simeq 1.3465$ and $r_{0.999}(\mathbb{Z}'_m) \simeq 1.2708$. Their values have been computed using very accurate approximations developed in Chapter 8, see Theorem 8.3.1. We will return to this example in Section 9.2.1.

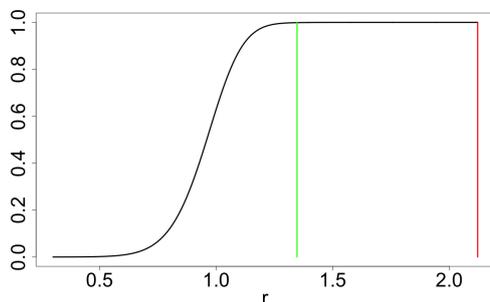


Figure 9.1: $C_d(\mathbb{Z}_n, r)$ with $r_{0.999}$ and r_1 : $d = 10$, \mathbb{Z}_n is a 2^{d-1} -factorial design with $n = 2^{d-1}$

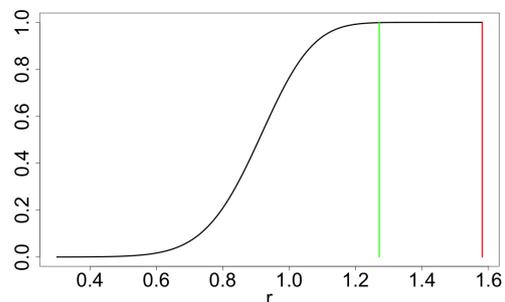


Figure 9.2: $C_d(\mathbb{Z}'_m, r)$ with $r_{0.999}$ and r_1 : $d = 10$, \mathbb{Z}_m is a 2^d -factorial design

Of course, for any $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ we can reach $C_d(\mathbb{Z}_n, r) = 1$ by means of increase of r . Likewise, for any given r we can reach $C_d(\mathbb{Z}_n, r) = 1$ by sending $n \rightarrow \infty$ (for any non-trivial design). However, we are not interested in very large values of n and try to get the coverage of the most part of the cube \mathcal{C}_d with the radius r as small as possible. We will keep in mind the following typical values of d and n which we will use for illustrating our results: $d = 5, 10, 20, 50$; $n = 2^k$ with $k = 6, \dots, 11$ (the author has chosen n as a power of 2 since this a favourable number for Sobol's sequence (Design 3) as well as Design 4 defined in Section 9.8).

¹For simplicity of notation, vectors in \mathbb{R}^d are represented as rows.

The structure of the rest of the chapter is as follows. In Section 9.2 we discuss the concept of weak covering in more detail and introduce three generic designs which we will concentrate our attention on. In Sections 9.3, 9.4 and 9.5 we derive approximations for the expected volume of intersection the cube $[-1, 1]^d$ with n balls centred at the points of these designs. In Section 9.6, we provide numerical results showing that the developed approximations are very accurate. In Section 9.7, we derive approximations for the mean squared quantization error for chosen families of designs and numerically demonstrate that the developed approximations are very accurate. In Section 9.8, we numerically compare covering and quantization properties of different designs including scaled Sobol's sequence and a family of very efficient designs defined only for very specific values of n . In Section 9.9 we numerically investigating the importance of the effect of scaling points away from the boundary (we call it δ -effect) for covering and quantization in a d -dimensional simplex. In Appendix, Section 9.10, we formulate a simple but important lemma about the distribution and moments of a certain random variable.

The main theoretical contributions of this chapter are:

- derivation of accurate approximations (9.3.8) and (9.4.3) for the probability $P_{U,\delta,\alpha,r}$ defined in (9.3.1);
- derivation of accurate approximations (9.3.10), (9.4.5) and (9.5.2) for the expected volume of intersection of the cube $[-1, 1]^d$ with n balls centred at the points of the selected designs;
- derivation of accurate approximations (9.7.4), (9.7.6) and (9.7.7) for the mean squared quantization error for the selected designs.

The author has performed a large-scale numerical study and provided a number of figures and tables. The following are the key messages containing in these figures and tables.

- Figures 9.1–9.2: weak covering could be much more practically useful than the full covering;
- Figures 9.3–9.14: developed approximations for the probability $P_{U,\delta,\alpha,r}$ defined in (9.3.1) are very accurate;
- Figures 9.15–9.28: (a) developed approximations for $C_d(\mathbb{Z}_n, r)$ are very accurate, (b) there is a very strong δ -effect for all three types of designs, and (c) this δ -effect gets stronger as d increases;
- Tables 9.1 and 9.2 and Figures 9.29–9.30: smaller values of α are beneficial in Design 1 but Design 2 (where $\alpha = 0$) becomes inefficient when n gets close to 2^d ;
- Figures 9.31–9.44: developed approximations for the quantization error are very accurate and there is a very strong δ -effect for all three types of designs used for quantization;

- Tables 9.3–9.4 and Figures 9.45–9.46: (a) Designs 2a and especially 2b provide very high quality coverage for suitable n , (b) properly δ -tuned deterministic non-nested Design 4 provides superior covering, (c) coverage properties of δ -tuned low-discrepancy sequences are much better than of the original low-discrepancy sequences, and (d) coverage properties of unadjusted low-discrepancy sequences is very low, if dimension d is not small;
- Tables 9.5 and 9.6, Figures 9.47 and 9.48: very similar conclusions to the above but made with respect to the quantization error;
- Figures 9.51–9.62: the δ -effect for covering and quantization schemes in a simplex is definitely present (this effect is more apparent in quantization) but it is much weaker than in a cube.

9.2 Weak covering

In this section, we consider the problem of weak covering defined and discussed in Section 9.1. The main characteristic of interest will be $C_d(\mathbb{Z}_n, r)$, the proportion of the cube covered by the union of balls $\mathcal{B}_d(\mathbb{Z}_n, r)$; it is defined in (9.1.1). We start the section with short discussion on comparison of designs based on their covering properties.

9.2.1 Comparison of designs from the view-point of weak covering

Two different designs will be differentiated in terms of covering performance as follows. Fix d and let \mathbb{Z}_n and \mathbb{Z}'_n be two n -point designs. For $(1 - \gamma)$ -covering with $\gamma \geq 0$, if $C_d(\mathbb{Z}_n, r) = C_d(\mathbb{Z}'_n, r') = 1 - \gamma$ and $r < r'$, then the design \mathbb{Z}_n provides a more efficient $(1 - \gamma)$ -covering and is therefore preferable. Moreover, the natural scaling for the radius is $r_n = n^{1/d}r$ and therefore we can compare an n -point design \mathbb{Z}_n with an m -point design \mathbb{Z}'_m as follows: if $C_d(\mathbb{Z}_n, r) = C_d(\mathbb{Z}'_m, r') = 1 - \gamma$ and $n^{1/d}r < m^{1/d}r'$, then we say that the design \mathbb{Z}_n provides a more efficient $(1 - \gamma)$ -covering than the design \mathbb{Z}'_m .

As an example, consider the designs used for plotting Figures 9.1 and 9.2 in Section 9.1: \mathbb{Z}_n with $n = 2^{d-1}$ and \mathbb{Z}'_m with $m = 2^d$. For the full covering, we have for any d :

$$n^{1/d}r_1(\mathbb{Z}_n) = 2^{-1/d}\sqrt{d+8} > \sqrt{d} = r_1(\mathbb{Z}'_m)m^{1/d}$$

so that the design \mathbb{Z}'_m is better than \mathbb{Z}_n for the full covering for any d and the difference between normalized covering radii is quite significant. For example, for $d = 10$ we have

$$n^{1/d}r_1(\mathbb{Z}_n) \simeq 3.9585 \quad \text{and} \quad r_1(\mathbb{Z}'_m)m^{1/d} \simeq 3.1623$$

For 0.999-covering, however, the situation is reverse, at least for $d = 10$, where we have:

$$n^{1/d}r_{0.999}(\mathbb{Z}_n) \simeq 2.5126 < 2.5416 \simeq r_1(\mathbb{Z}'_m)m^{1/d}$$

and therefore the design \mathbb{Z}_n is better for 0.999-covering than the design \mathbb{Z}'_m for $d = 10$.

9.2.2 Reduction to the probability of covering a point by one ball

In the designs \mathbb{Z}_n , which are of most interest to us, the points $Z_j \in \mathbb{Z}_n$ are i.i.d. random vectors in \mathbb{R}^d with a specified distribution. Let us show that for these designs, we can reduce computation of $C_d(\mathbb{Z}_n, r)$ to the probability of covering $[-1, 1]^d$ by one ball. This is performed in the same manner as Chapter 7, see Section 7.3.2

Let Z_1, \dots, Z_n be i.i.d. random vectors in \mathbb{R}^d and $\mathcal{B}_d(\mathbb{Z}_n, r)$ be as defined in (9.1.2). Then, for given $U = (u_1, \dots, u_d) \in \mathbb{R}^d$ (we will later assume U is uniformly distributed in $[-1, 1]^d$):

$$\begin{aligned} \mathbb{P}\{U \in \mathcal{B}_d(\mathbb{Z}_n, r)\} &= 1 - \prod_{j=1}^n \mathbb{P}\{U \notin \mathcal{B}_d(Z_j, r)\} \\ &= 1 - \prod_{j=1}^n (1 - \mathbb{P}\{U \in \mathcal{B}_d(Z_j, r)\}) \\ &= 1 - \left(1 - \mathbb{P}_Z\{\|U - Z\| \leq r\}\right)^n. \end{aligned} \quad (9.2.1)$$

$C_d(\mathbb{Z}_n, r)$, defined in (9.1.1), is simply

$$C_d(\mathbb{Z}_n, r) = \mathbb{E}_U \mathbb{P}\{U \in \mathcal{B}_d(\mathbb{Z}_n, r)\}, \quad (9.2.2)$$

where the expectation is taken with respect to the uniformly distributed $U \in [-1, 1]^d$. For numerical convenience, we shall simplify the expression (9.2.1) by using the approximation

$$(1 - t)^n \simeq e^{-nt}, \quad (9.2.3)$$

where $t = \mathbb{P}_Z\{\|U - Z\| \leq r\}$. This approximation is very accurate for small values of t and moderate values of nt , which is always the case of our interest. Combining (9.2.1), (9.2.2) and (9.2.3), we obtain the approximation

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \mathbb{E}_U \exp(-n \cdot \mathbb{P}_Z\{\|U - Z\| \leq r\}). \quad (9.2.4)$$

In the next section we will formulate three schemes that will be of theoretical interest in this chapter. For each scheme and hence different distribution of Z , we shall derive accurate approximations for $\mathbb{P}_Z\{\|U - Z\| \leq r\}$ and therefore, using (9.2.4), for $C_d(\mathbb{Z}_n, r)$.

9.2.3 Designs of theoretical interest

The three designs that will be the focus of theoretical investigation in this chapter are:

Design 1. $Z_1, \dots, Z_n \in \mathbb{Z}_n$ are i.i.d. random vectors on $[-\delta, \delta]^d$ with independent components distributed according to the following $\text{Beta}_\delta(\alpha, \alpha)$ distribution with density:

$$p_{\alpha, \delta}(t) = \frac{(2\delta)^{1-2\alpha}}{\text{Beta}(\alpha, \alpha)} [\delta^2 - t^2]^{\alpha-1}, \quad -\delta < t < \delta, \text{ for } \alpha > 0 \text{ and } 0 \leq \delta \leq 1. \quad (9.2.5)$$

Design 2a. $Z_1, \dots, Z_n \in \mathbb{Z}_n$ are i.i.d. random vectors obtained by sampling with replacement from the vertices of the cube $[-\delta, \delta]^d$.

Design 2b. $Z_1, \dots, Z_n \in \mathbb{Z}_n$ are random vectors obtained by sampling without replacement from the vertices of the cube $[-\delta, \delta]^d$.

All three designs above are nested so that $\mathbb{Z}_n \subset \mathbb{Z}_{n+1}$ for all eligible n . Designs 1 and 2a are defined for all $n = 1, 2, \dots$ whereas Design 2b is defined for $n = 1, 2, \dots, 2^d$. The appealing property of any design whose points Z_i are i.i.d. is the possibility of using (9.2.1); this is the case of Designs 1 and 2a. For Design 2b, we will need to make some adjustments, see Section 9.5. Design 1. has previously been studied in Chapter 7, where we called it Scheme 4, and was shown to have some appealing covering properties along with its nested nature.

In the case of $\alpha = 1$ in Design 1, the distribution $\text{Beta}_\delta(\alpha, \alpha)$ becomes uniform on $[-\delta, \delta]^d$. This case has been comprehensively studied in Chapter 7 with a number of approximations for $C_d(\mathbb{Z}_n, r)$ being developed. The approximations developed in Section 9.3 are generalizations of the approximations of Chapter 7. Numerical results of Chapter 7 indicated that Beta-distribution with $\alpha < 1$ provides more efficient covering schemes; this explains the importance of the approximations of Section 9.3. Design 2a is the limiting form of Design 1 as $\alpha \rightarrow 0$. Theoretical approximations developed below for $C_d(\mathbb{Z}_n, r)$ for Design 2a are, however, more precise than the limiting cases of approximations obtained for $C_d(\mathbb{Z}_n, r)$ in case of Design 1. For numerical comparison, in Section 9.6 we shall also consider several other designs.

9.3 Approximation of $C_d(\mathbb{Z}_n, r)$ for Design 1

As a result of (9.2.4), our main quantity of interest in this section will be the probability

$$P_{U,\delta,\alpha,r} := \mathbb{P}_Z \{ \|U - Z\| \leq r \} = \mathbb{P}_Z \{ \|U - Z\|^2 \leq r^2 \} = \mathbb{P} \left\{ \sum_{j=1}^d (u_j - z_j)^2 \leq r^2 \right\} \quad (9.3.1)$$

in the case when Z has the Beta-distribution with density (9.2.5). We shall develop a simple approximation based on the Central Limit Theorem (CLT) and then subsequently refine it using the general expansion in the CLT for sums of independent non-identical r.v.

9.3.1 Normal approximation for $P_{U,\delta,\alpha,r}$

Let $\eta_{u,\delta,\alpha} = (z - u)^2$, where z has density (9.2.5). In view of Lemma 9.10.1, the r.v. $\eta_{u,\delta,\alpha}$ is concentrated on the interval $[(\max(0, \delta - |u|))^2, (\delta + |u|)^2]$ and its first three

central moments are:

$$\mu_u^{(1)} = \mathbb{E}\eta_{u,\delta,\alpha} = u^2 + \frac{\delta^2}{2\alpha + 1}, \quad (9.3.2)$$

$$\mu_u^{(2)} = \text{var}(\eta_{u,\delta,\alpha}) = \frac{4\delta^2}{2\alpha + 1} \left[u^2 + \frac{\delta^2\alpha}{(2\alpha + 1)(2\alpha + 3)} \right], \quad (9.3.3)$$

$$\mu_u^{(3)} = \mathbb{E} \left[\eta_{u,\delta,\alpha} - \mu_u^{(1)} \right]^3 = \frac{48\alpha\delta^4}{(2\alpha + 1)^2(2\alpha + 3)} \left[u^2 + \frac{\delta^2(2\alpha - 1)}{3(2\alpha + 5)(2\alpha + 1)} \right]. \quad (9.3.4)$$

For a given $U = (u_1, \dots, u_d) \in \mathbb{R}^d$, consider the r.v.

$$\|U - Z\|^2 = \sum_{i=1}^d \eta_{u_i,\delta,\alpha} = \sum_{j=1}^d (u_j - z_j)^2,$$

where we assume that $Z = (z_1, \dots, z_d)$ is a random vector with i.i.d. components z_i with density (9.2.5). From (9.3.2), its mean is

$$\mu = \mu_{d,\delta,\alpha,U} := \mathbb{E}\|U - Z\|^2 = \|U\|^2 + \frac{d\delta^2}{2\alpha + 1}.$$

Using independence of z_1, \dots, z_d and (9.3.3), we obtain

$$\sigma_{d,\delta,\alpha,U}^2 := \text{var}(\|U - Z\|^2) = \frac{4\delta^2}{2\alpha + 1} \left[\|U\|^2 + \frac{d\delta^2\alpha}{(2\alpha + 1)(2\alpha + 3)} \right],$$

and from independence of z_1, \dots, z_d and (9.3.4) we get

$$\begin{aligned} \mu_{d,\delta,\alpha,U}^{(3)} &:= \mathbb{E} [\|U - Z\|^2 - \mu]^3 \\ &= \sum_{j=1}^d \mu_{u_j}^{(3)} = \frac{48\alpha\delta^4}{(2\alpha + 1)^2(2\alpha + 3)} \left[\|U\|^2 + \frac{d\delta^2(2\alpha - 1)}{3(2\alpha + 5)(2\alpha + 1)} \right]. \end{aligned}$$

If d is large enough then the conditions of the CLT for $\|U - Z\|^2$ are approximately met and the distribution of $\|U - Z\|^2$ is approximately normal with mean $\mu_{d,\delta,\alpha,U}$ and variance $\sigma_{d,\delta,\alpha,U}^2$. That is, we can approximate the probability $P_{U,\delta,\alpha,r} = \mathbb{P}_Z \{\|U - Z\| \leq r\}$ by

$$P_{U,\delta,\alpha,r} \cong \Phi \left(\frac{r^2 - \mu_{d,\delta,\alpha,U}}{\sigma_{d,\delta,\alpha,U}} \right), \quad (9.3.5)$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution:

$$\Phi(t) = \int_{-\infty}^t \varphi(v) dv \quad \text{with} \quad \varphi(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2}.$$

The approximation (9.3.5) has acceptable accuracy if the probability $P_{U,\delta,\alpha,r}$ is not very small; for example, it falls inside a 2σ -confidence interval generated by the standard normal distribution. In the next section, we improve approximations (9.3.5) by using an Edgeworth-type expansion in the CLT for sums of independent non-identically distributed r.v.

9.3.2 Refined approximation for $P_{U,\delta,\alpha,r}$

From the discussion of Section 7.2.4 in Chapter 7, which is based around the general expansion in the central limit theorem for sums of independent non-identical r.v. derived by V. Petrov, we obtain (by using the first term in this expansion) the following approximation for the distribution function of $\|U - Z\|^2$:

$$\mathbb{P}\left(\frac{\|U - Z\|^2 - \mu_{d,\delta,\alpha,U}}{\sigma_{d,\delta,\alpha,U}} \leq x\right) \cong \Phi(x) + \frac{\mu_{d,\delta,\alpha,U}^{(3)}}{6\sigma_{d,\delta,\alpha,U}^3}(1 - x^2)\varphi(x), \quad (9.3.6)$$

leading to the following improved form of (9.3.5):

$$P_{U,\delta,\alpha,r} \cong \Phi(t) + \frac{\alpha\delta \left[\|U\|^2 + \frac{d\delta^2(2\alpha-1)}{3(2\alpha+5)(2\alpha+1)} \right]}{(2\alpha+3)(2\alpha+1)^{1/2} \left[\|U\|^2 + \frac{d\delta^2\alpha}{(2\alpha+1)(2\alpha+3)} \right]^{3/2}} (1-t^2)\varphi(t), \quad (9.3.7)$$

where

$$t := \frac{r^2 - \mu_{d,\delta,\alpha,U}}{\sigma_{d,\delta,\alpha,U}} = \frac{\sqrt{2\alpha+1}(r^2 - \|U\|^2 - \frac{d\delta^2}{2\alpha+1})}{2\delta\sqrt{\|U\|^2 + \frac{d\delta^2\alpha}{(2\alpha+1)(2\alpha+3)}}}.$$

For $\alpha = 1$, we obtain

$$P_{U,\delta,\alpha,r} \cong \Phi(t) + \frac{\delta \left[\|U\|^2 + d\delta^2/63 \right]}{5\sqrt{3} \left[\|U\|^2 + d\delta^2/15 \right]^{3/2}} (1-t^2)\varphi(t)$$

with $t = \frac{\sqrt{3}(r^2 - \|U\|^2 - d\delta^2/3)}{2\delta\sqrt{\|U\|^2 + d\delta^2/15}},$

which coincides with formula (7.2.15) of Chapter 7 (with $\delta = 1$).

A very attractive feature of the approximations (9.3.5) and (9.3.7) is their dependence on U through $\|U\|$ only. We could have specialized for our case the next terms in Petrov's approximation but these terms no longer depend on $\|U\|$ only and hence the next terms are much more complicated. Moreover, adding one or two extra terms from Petrov's expansion to the approximation (9.3.7) does not fix the problem entirely for all U , δ , α and r . Instead, the author proposes a slight adjustment to the r.h.s of (9.3.7) to improve this approximation, especially for small dimensions. Specifically, the author suggests the approximation

$$P_{U,\delta,\alpha,r} \cong \Phi(t) + \tag{9.3.8}$$

$$c_{d,\alpha} \frac{\alpha\delta \left[\|U\|^2 + \frac{d\delta^2(2\alpha-1)}{3(2\alpha+5)(2\alpha+1)} \right]}{(2\alpha+3)(2\alpha+1)^{1/2} \left[\|U\|^2 + \frac{d\delta^2\alpha}{(2\alpha+1)(2\alpha+3)} \right]^{3/2}} (1-t^2)\varphi(t),$$

where $c_{d,\alpha} = 1 + 3/(ad)$.

Below, there are figures of two types. In Figures 9.3–9.4, we plot $P_{U,\delta,\alpha,r}$ over a wide range of r ensuring that values of $P_{U,\delta,\alpha,r}$ lie in the whole range $[0, 1]$. In Figures 9.5–9.8, we plot $P_{U,\delta,\alpha,r}$ over a much smaller range of r with $P_{U,\delta,\alpha,r}$ lying

roughly in the range $[0, 0.02]$. For the purpose of using formula (9.2.1), we need to assess the accuracy of all approximations for smaller values of $P_{U,\delta,\alpha,r}$ and hence the second type of plots are more useful. In these figures, the solid black line depicts $P_{U,\delta,\alpha,r}$ obtained via Monte Carlo methods with 50,000 iterations where for simplicity we have set $U = (1/2, 1/2, \dots, 1/2)$ and $\delta = 1/2$. Approximations (9.3.5) and (9.3.8) are depicted with a dotted blue and dash green line respectively. From numerous simulations and these figures, we can conclude the following. Whilst the basic normal approximation (9.3.5) seems adequate in the whole range of values of r , for particularly small probabilities, that we are most interested in, approximation (9.3.8) is much superior and appears to be very accurate for all values of α .

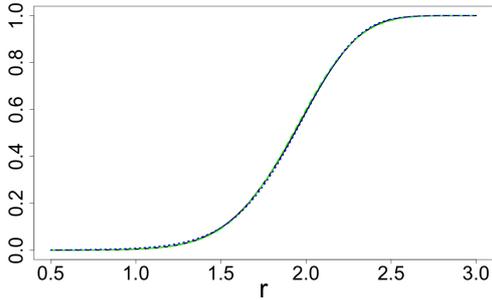


Figure 9.3: $P_{U,\delta,\alpha,r}$ and approximations: $d = 10$, $\alpha = 0.5$.

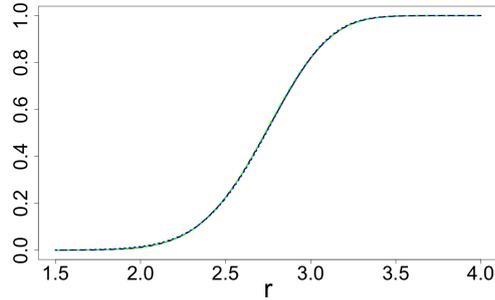


Figure 9.4: $P_{U,\delta,\alpha,r}$ and approximations: $d = 20$, $\alpha = 0.5$.

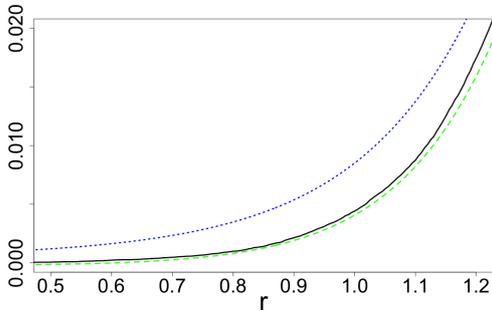


Figure 9.5: $P_{U,\delta,\alpha,r}$ and approximations: $d = 10$, $\alpha = 0.5$.

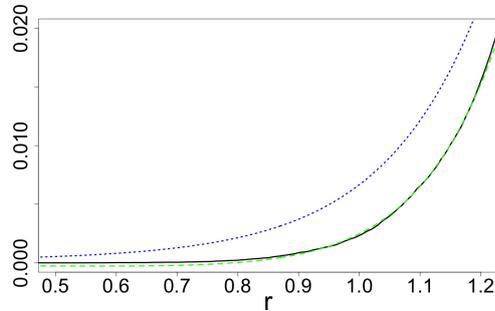


Figure 9.6: $P_{U,\delta,\alpha,r}$ and approximations: $d = 10$, $\alpha = 1$.

9.3.3 Approximation for $C_d(\mathbb{Z}_n, r)$ for Design 1

Consider now $C_d(\mathbb{Z}_n, r)$ for Design 1, as expressed via $P_{U,\delta,\alpha,r}$ in (9.2.4). As U is uniform on $[-1, 1]^d$, $\mathbb{E}\|U\|^2 = d/3$ and $\text{var}(\|U\|^2) = 4d/45$. Moreover, if d is large enough then $\|U\|^2 = \sum_{j=1}^d u_j^2$ is approximately normal.

We will combine the expressions (9.2.4) with approximations (9.3.5) and (9.3.8) as well as with the normal approximation for the distribution of $\|U\|^2$, to arrive at two final approximations for $C_d(\mathbb{Z}_n, r)$ that differ in complexity. If the original normal approximation (9.3.5) of $P_{U,\delta,\alpha,r}$ is used then we obtain:

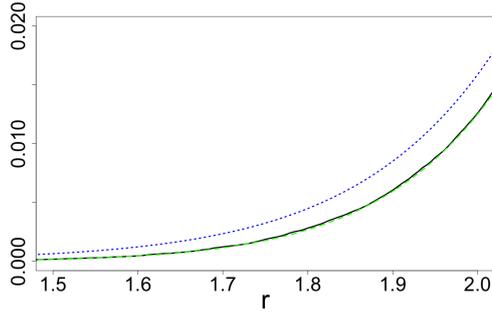


Figure 9.7: $P_{U,\delta,\alpha,r}$ and approximations: $d = 20$, $\alpha = 0.5$.

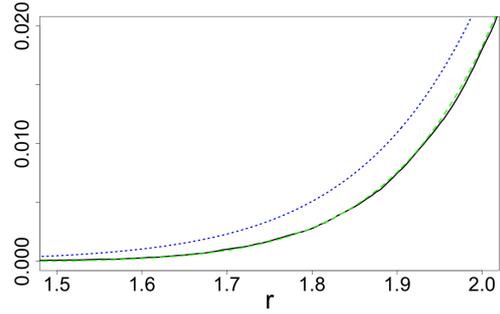


Figure 9.8: $P_{U,\delta,\alpha,r}$ and approximations: $d = 20$, $\alpha = 1$.

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_{1,\alpha}(s, r) \varphi(s) ds \quad (9.3.9)$$

with

$$\psi_{1,\alpha}(s, r) = \exp\{-n\Phi(c_{s,r})\}, \quad c_{s,r} = \frac{(2\alpha + 1)^{1/2} \left(r^2 - s' - \frac{d\delta^2}{2\alpha+1} \right)}{2\delta\sqrt{s' + \kappa}}, \quad s' = s\sqrt{\frac{4d}{45}} + d/3,$$

$$\kappa = \frac{d\delta^2\alpha}{(2\alpha+1)(2\alpha+3)}.$$

If the approximation (9.3.8) is used, we obtain:

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_{2,\alpha}(s, r) \varphi(s) ds, \quad (9.3.10)$$

with

$$\psi_{2,\alpha}(s, r) = \exp \left\{ -n \left(\Phi(c_s) + c_{d,\alpha} \frac{\alpha\delta \left[s' + \frac{d\delta^2(2\alpha-1)}{3(2\alpha+5)(2\alpha+1)} \right]}{(2\alpha+3)(2\alpha+1)^{1/2} [s' + \kappa]^{3/2}} (1 - c_{s,r}^2) \varphi(c_{s,r}) \right) \right\}.$$

For $\alpha = 1$, we get

$$\psi_{2,1}(s, r) = \exp \left\{ -n \left(\Phi(c_{s,r}) + c_{d,\alpha} \frac{\delta \left[s' + \frac{d\delta^2}{63} \right]}{5\sqrt{3} \left[s' + \frac{d\delta^2}{15} \right]^{3/2}} (1 - c_{s,r}^2) \varphi(c_{s,r}) \right) \right\} \quad (9.3.11)$$

and the approximation (9.3.10) coincides with the approximation (7.3.8) in Chapter 7. The accuracy of approximations (9.3.9) and (9.3.10) will be assessed in Section 9.6.1.

9.4 Approximating $C_d(\mathbb{Z}_n, r)$ for Design 2a

Our main quantity of interest in this section will be the probability $P_{U,\delta,0,r}$ defined in (9.3.1) in the case where components z_i of the vector $Z = (z_1, \dots, z_d) \in \mathbb{R}^d$ are i.i.d.r.v with $\Pr(z_i = \delta) = \Pr(z_i = -\delta) = 1/2$; this is a limiting case of $P_{U,\delta,\alpha,r}$ as $\alpha \rightarrow 0$.

9.4.1 Normal approximation for $P_{U,\delta,0,r}$

Using the same approach that led to approximation (9.3.5) in Section 9.3.1, the initial normal approximation for $P_{U,\delta,0,r}$ is:

$$P_{U,\delta,0,r} \cong \Phi\left(\frac{r^2 - \mu_{d,\delta,U}}{\sigma_{d,\delta,U}}\right), \quad (9.4.1)$$

where, from Lemma 9.10.1, we have

$$\mu_{d,\delta,U} = \|U\|^2 + d\delta^2 \quad \text{and} \quad \sigma_{d,\delta,U}^2 = 4\delta^2\|U\|^2.$$

9.4.2 Refined approximation for $P_{U,\delta,0,r}$

From (9.10.2), we have $\mu_{d,\delta,\alpha,U}^{(3)} = 0$ and therefore the last term in the rhs of (9.3.6) with $\alpha = 0$ is no longer present. By taking an additional term in the general expansion, see V. Petrov in Section 5.6 in [89], we obtain the following approximation for the distribution function of $\|U - Z\|^2$:

$$\mathbb{P}\left(\frac{\|U - Z\|^2 - \mu_{d,\delta,U}}{\sigma_{d,\delta,U}} \leq x\right) \cong \Phi(x) - (x^3 - 3x) \frac{\kappa_{d,\delta,0,U}^{(4)}}{24\sigma_{d,\delta,0,U}^4} \varphi(x), \quad (9.4.2)$$

where $\kappa_{d,\delta,0,U}^{(4)}$ is the sum of d fourth cumulants of the centred r.v. $(z - u)^2$, where z is concentrated at two points $\pm\delta$ with $\Pr(z = \pm\delta) = 1/2$. From (9.10.2),

$$\kappa_{d,\delta,0,U}^{(4)} := \sum_{j=1}^d (\mu_{u_j}^{(4)} - 3[\mu_{u_j}^{(2)}]^2) = -32\delta^4 \sum_{i=1}^d u_i^4.$$

Unlike (9.3.6), the rhs of (9.4.2) does not depend solely on $\|U\|^2$. However, the quantities $\|U\|^2$ and $\sum_{i=1}^d u_i^4$ are strongly correlated; one can show that for all d

$$\text{corr}\left(\|U\|^2, \sum_{i=1}^d u_i^4\right) = \frac{3\sqrt{5}}{7} \cong 0.958.$$

This suggests (by rounding the correlation above to 1) the following approximation:

$$\sum_{i=1}^d u_i^4 \cong \frac{4\sqrt{d}}{15} \left(\frac{\|U\|^2 - d/3}{\sqrt{\frac{4d}{45}}}\right) + \frac{d}{5}.$$

With this approximation, the rhs of (9.4.2) depends only on $\|U\|^2$. As a result, the following refined form of (9.4.1) is:

$$P_{U,\delta,0,r} \cong \Phi(t) + (t^3 - 3t) \frac{2(\|U\|^2 - d/3)/\sqrt{5} + d/5}{12\|U\|^4} \varphi(t),$$

where

$$t := \frac{r^2 - \mu_{d,\delta,0,U}}{\sigma_{d,\delta,0,U}} = \frac{(r^2 - \|U\|^2 - d\delta^2)}{2\delta\|U\|}.$$

Similarly to approximation (9.3.8), a slight adjustment to the r.h.s of the approximation above is recommended:

$$P_{U,\delta,0,r} \cong \Phi(t) + \left(1 + \frac{3}{d}\right) (t^3 - 3t) \frac{2(\|U\|^2 - d/3)/\sqrt{5} + d/5}{12\|U\|^4} \varphi(t). \quad (9.4.3)$$

In the same style as at the end of Section 9.3.2, below there are figures of two types. In Figures 9.9–9.10, we plot $P_{U,\delta,0,r}$ over a wide range of r ensuring that values of $P_{U,\delta,0,r}$ lie in the range $[0, 1]$. In Figures 9.11–9.14, we plot $P_{U,\delta,0,r}$ over a much smaller range of r with $P_{U,\delta,0,r}$ lying in the range $[0, 0.02]$. In these figures, the solid black line depicts $P_{U,\delta,\alpha,r}$ obtained via Monte Carlo methods with 50,000 repetitions where we have set $\delta = 1/2$ and U is a point sampled uniformly on $[-1, 1]^d$; for reproducibility, in the caption of each figure the random seed used in R is provided. Approximations (9.4.1) and (9.4.3) are depicted with a dotted blue and dash green line respectively. From these figures, we can conclude the same outcome as in Section 9.3.2. Whilst the approximation (9.4.1) is rather good overall, for small probabilities the approximation (9.4.3) is much superior and is very accurate. Note that since random vectors Z_j are taking values on a finite set, which is the set of points $(\pm\delta, \dots, \pm\delta)$, the probability $P_{U,\delta,0,r}$ considered as a function of r , is a piece-wise constant function.

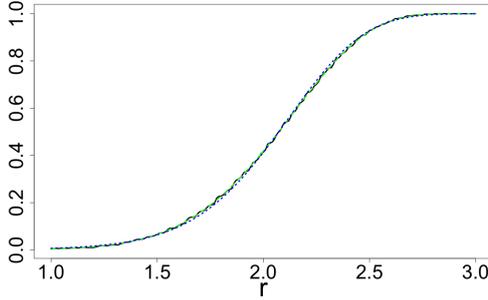


Figure 9.9: $P_{U,\delta,0,r}$ and approximations: $d = 10$, $seed = 10$.

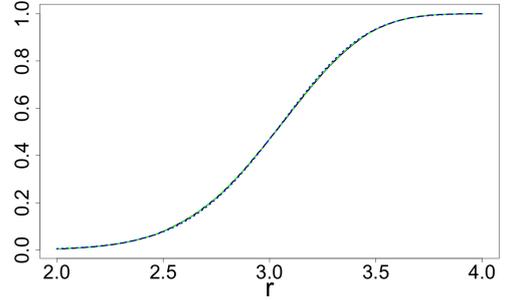


Figure 9.10: $P_{U,\delta,0,r}$ and approximations: $d = 20$, $seed = 10$.

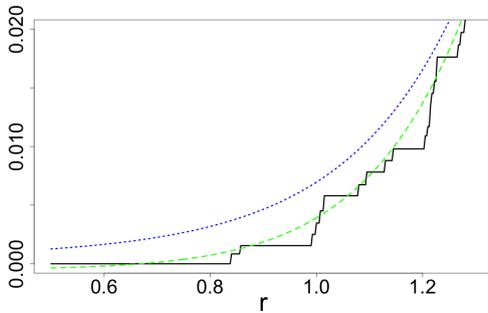


Figure 9.11: $P_{U,\delta,0,r}$ and approximations: $d = 10$, $seed = 10$.

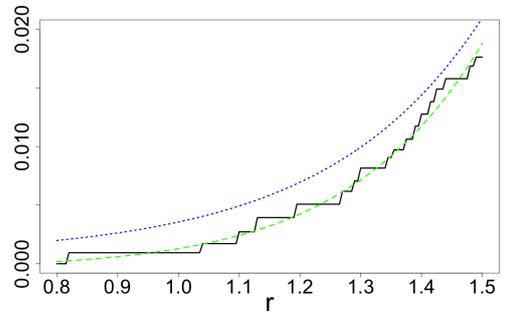


Figure 9.12: $P_{U,\delta,0,r}$ and approximations: $d = 10$, $seed = 15$.

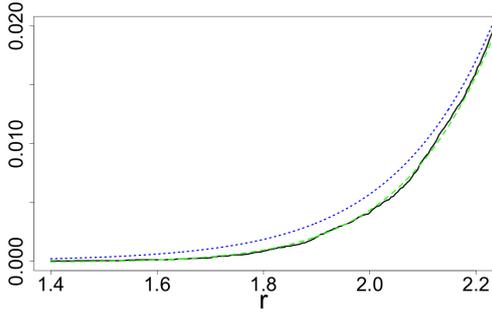


Figure 9.13: $P_{U,\delta,0,r}$ and approximations: $d = 20$, $seed = 10$.

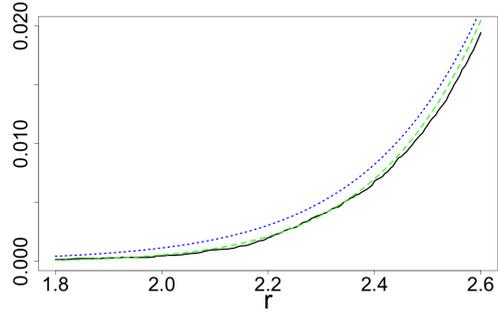


Figure 9.14: $P_{U,\delta,0,r}$ and approximations: $d = 20$, $seed = 15$.

9.4.3 Approximation for $C_d(\mathbb{Z}_n, r)$

Consider now $C_d(\mathbb{Z}_n, r)$ for Design 2a, as expressed via $P_{U,\delta,\alpha,r}$ in (9.2.4). Using the normal approximation for $\|U\|^2$ as made in the beginning of Section 9.3.3, we will combine the expressions (9.2.4) with approximations (9.4.1) and (9.4.3) to arrive at two approximations for $C_d(\mathbb{Z}_n, r)$ that differ in complexity.

If the original normal approximation (9.4.1) of $P_{U,\delta,0,r}$ is used then we obtain:

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_{3,n}(s, r) \varphi(s) ds, \quad (9.4.4)$$

with

$$\psi_{3,n}(s) = \exp \{-n\Phi(c_{s,r})\}, \quad c_{s,r} = \frac{(r^2 - s' - d\delta^2)}{2\delta\sqrt{s'}}, \quad s' = s\sqrt{\frac{4d}{45}} + d/3.$$

If the approximation (9.4.3) is used, we obtain:

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_{4,n}(s, r) \varphi(s) ds, \quad (9.4.5)$$

with

$$\psi_{4,n}(s, r) = \exp \left\{ -n \left(\Phi(c_{s,r}) + \left(1 + \frac{3}{d}\right) (c_{s,r}^3 - 3c_{s,r}) \frac{2(s' - d/3)/\sqrt{5} + d/5}{12(s')^2} \varphi(c_{s,r}) \right) \right\} \quad (9.4.6)$$

and

$$c_{s,r} = \frac{(r^2 - s' - d\delta^2)}{2\delta\sqrt{s'}}, \quad s' = s\sqrt{\frac{4d}{45}} + d/3.$$

The accuracy of approximations (9.4.4) and (9.4.5) will be assessed in Section 9.6.1.

9.5 Approximating $C_d(\mathbb{Z}_n, r)$ for Design 2b

Designs whose points Z_i have been sampled from a finite discrete set without replacement have dependence, for example Design 2b, and therefore formula (9.2.1) cannot be used.

In this section, we suggest a way of modifying the approximations developed in Section 9.4 for Design 2a. This will amount to approximating sampling without replacement by a suitable sampling with replacement.

9.5.1 Establishing a connection between sampling with and without replacement: general case

Let \mathcal{S} be a discrete set with k distinct elements, where k is reasonably large. In case of Design 2b, the set \mathcal{S} consists of $k = 2^d$ vertices of the cube $[-\delta, \delta]^d$. Let $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$ denote an n -point design whose points Z_i have been sampled without replacement from \mathcal{S} ; $n < k$. Also, let $\mathbb{Z}'_m = \{Z'_1, \dots, Z'_m\}$ denote an associated m -point design whose points Z'_i are sampled with replacement from the same discrete set \mathcal{S} ; Z'_1, \dots, Z'_m are i.i.d. random vectors with values in \mathcal{S} . Our aim in this section is to establish an approximate correspondence between n and m .

When sampling m times with replacement, denote by X_i the number of times the i^{th} element of \mathcal{S} appears. Then the vector (X_1, X_2, \dots, X_k) has the multinomial distribution with number of trials m and event probabilities $(1/k, 1/k, \dots, 1/k)$ with each individual X_i having the Binomial distribution $\text{Binomial}(m, 1/k)$. Since $\text{corr}(X_i, X_j) = -1/k^2$ when $i \neq j$, for large k the correlation between random variables X_1, X_2, \dots, X_k is very small and will be neglected. Introduce the random variables:

$$Y_i = \begin{cases} 1, & \text{if } X_i = 0 \\ 0, & \text{if } X_i > 0. \end{cases}$$

Then the random variable $N_0 = \sum_{i=1}^k Y_i$ represents the number of elements of \mathcal{S} not selected. Given the weak correlation between X_i , we approximately have $N_0 \sim \text{Binomial}(k, P(X_1 = 0))$. Using the fact $P(X_1 = 0) = (1 - 1/k)^m$, the expected number of unselected elements when sampling with replacement is approximately $\mathbb{E}N_0 \cong k(1 - 1/k)^m$. Since, when sampling without replacement from \mathcal{S} we have chosen $N_0 = k - n$ elements, to choose the value of m we equate $\mathbb{E}N_0$ to $k - n$. By solving the equation

$$k - n = k \left(1 - \frac{1}{k}\right)^m \quad [\cong \mathbb{E}N_0]$$

for m we obtain

$$m = \frac{\log(k - n) - \log(k)}{\log(k - 1) - \log(k)}. \quad (9.5.1)$$

9.5.2 Approximation of $C_d(\mathbb{Z}_n, r)$ for Design 2b.

Consider now $C_d(\mathbb{Z}_n, r)$ for Design 2b. By applying the approximation developed in the previous section, the quantity $C_d(\mathbb{Z}_n, r)$ can be approximated by $C_d(\mathbb{Z}_m, r)$ for Design 2a with m given in (9.5.1):

Approximation of $C_d(\mathbb{Z}_n, r)$ for Design 2b. *We approximate it by $C_d(\mathbb{Z}_m, r)$ where m is given in (9.5.1) and $C_d(\mathbb{Z}_m, r)$ is approximated by (9.4.5) with n substituted by m from (9.5.1).*

Specifying this, we obtain:

$$C_d(\mathbb{Z}_n, r) \simeq 1 - \int_{-\infty}^{\infty} \psi_{4,m}(s, r) \varphi(s) ds, \quad (9.5.2)$$

where

$$m = m_{n,d} = \frac{\log(2^d - n) - d \log(2)}{\log(2^d - 1) - d \log(2)} \quad (9.5.3)$$

and the function $\psi_{4,\cdot}(\cdot, r)$ is defined in (9.4.6). The accuracy of the approximation (9.5.2) will be assessed in Section 9.6.1.

9.6 Numerical study

9.6.1 Assessing accuracy of approximations of $C_d(\mathbb{Z}_n, r)$ and studying their dependence on δ

In this section, we present the results of a large-scale numerical study assessing the accuracy of approximations (9.3.9), (9.3.10), (9.4.4), (9.4.5) and (9.5.2). In Figures 9.15–9.28, by using a solid black line we depict $C_d(\mathbb{Z}_n, r)$ obtained by Monte Carlo methods with 50,000 run, where the value of r has been chosen such that the maximum coverage across δ is approximately 0.9. In Figures 9.15–9.20, dealing with Design 1, approximations (9.3.9) and (9.3.10) are depicted with a dotted blue and dashed green lines respectively. In Figures 9.21–9.24 (Design 2a) approximations (9.4.4) and (9.4.5) are illustrated with a dotted blue and dashed green lines respectively. In Figures 9.25–9.28 (Design 2b) the dashed green line depicts approximation (9.5.2). From these figures, we can draw the following conclusions.

- Approximations (9.3.10) and (9.4.5) are very accurate across all values of δ and α . This is particularly evident for $d = 20, 50$.
- Approximations (9.3.9) and (9.4.4) are accurate only for very large values of d , like $d = 50$.
- Approximation (9.4.5) is generally accurate. For δ close to one (for such values of δ the covering is very poor) and n close to 2^d this approximation begins to worsen, see Figures 9.26 and 9.28.
- A sensible choice of δ can dramatically increase the coverage proportion $C_d(\mathbb{Z}_n, r)$. This effect, which we call ‘ δ -effect’, is evident in all figures and is very important. It gets much stronger as d increases.

9.6.2 Comparison across α

In Table 9.1, for Design 2a and Design 1 with $\alpha = 0.5, 1, 1.5$ we present the smallest values of r required to achieve the 0.9-coverage on average. For these schemes, the value inside the brackets shows the average value of δ required to obtain this 0.9-coverage. Design 2b is not used as d is too small (for this design, we must have

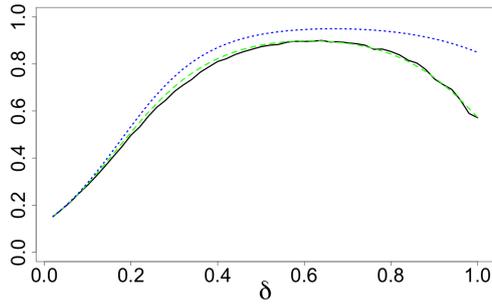


Figure 9.15: Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 10, \alpha = 0.5, n = 128$.

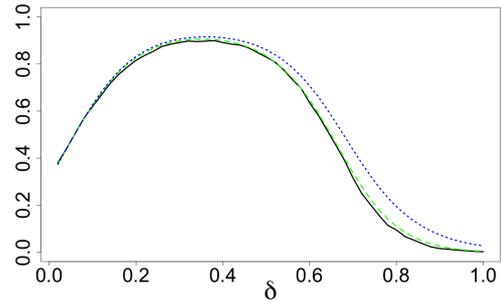


Figure 9.16: Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20, \alpha = 0.1, n = 128$.

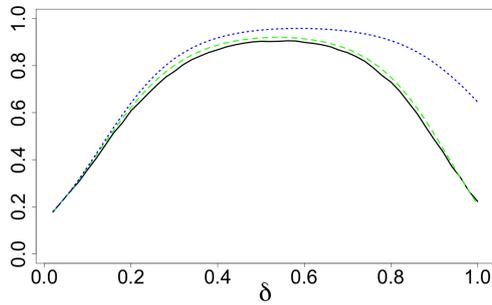


Figure 9.17: Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20, \alpha = 0.5, n = 512$.

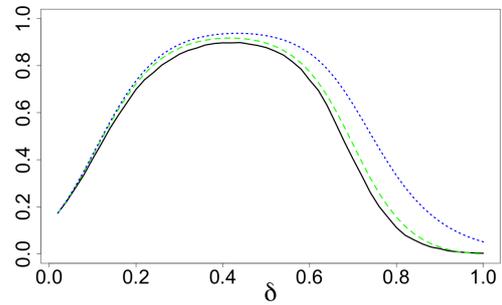


Figure 9.18: Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20, \alpha = 0.1, n = 512$.

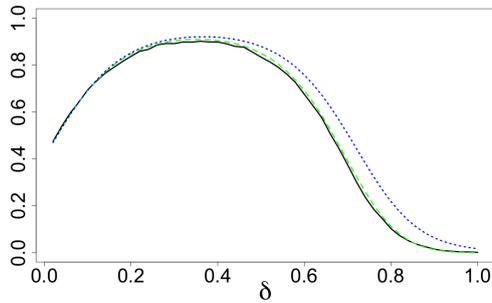


Figure 9.19: Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 50, \alpha = 0.5, n = 512$.

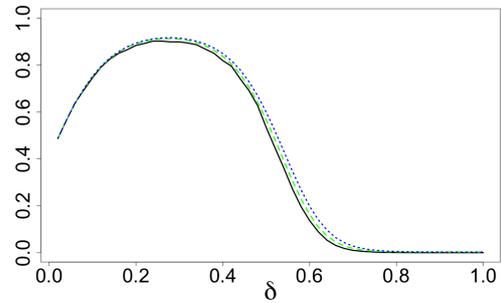


Figure 9.20: Design 1: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 50, \alpha = 0.1, n = 512$.

$n < 2^d$ and in these cases Design 2b provides better coverings than the other designs considered).

From Tables 9.1 and 9.2 we can make the following conclusions:

- For small n ($n < 2^d$ or $n \simeq 2^d$), Design 2a provides a more efficient covering than other three other schemes and hence smaller values of α are better.
- For $n > 2^d$, Design 2a begins to become impractical since a large proportion

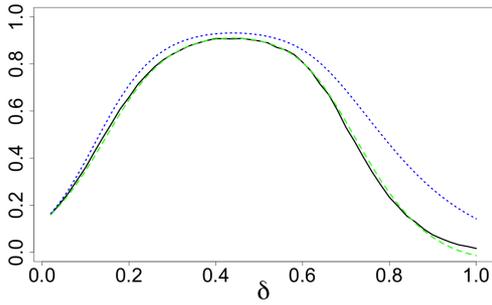


Figure 9.21: Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 10, \alpha = 0, n = 128$.

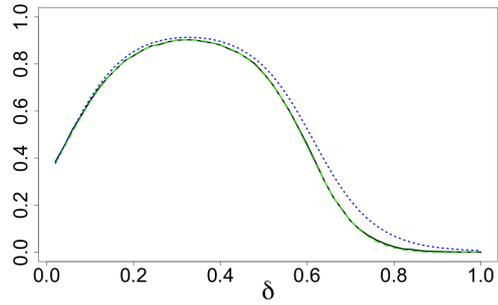


Figure 9.22: Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20, \alpha = 0, n = 128$.

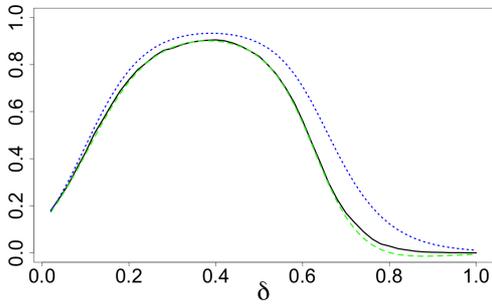


Figure 9.23: Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 20, \alpha = 0, n = 512$.

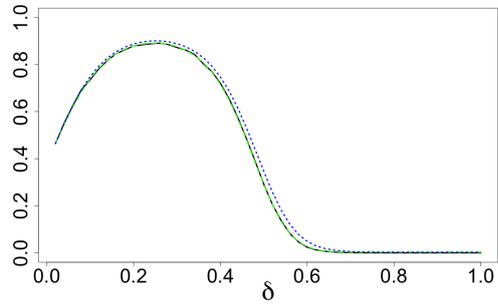


Figure 9.24: Design 2a: $C_d(\mathbb{Z}_n, r)$ and approximations; $d = 50, \alpha = 0, n = 512$.

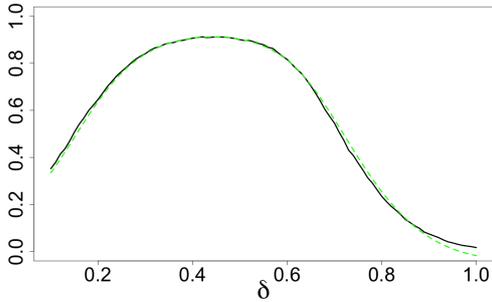


Figure 9.25: Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 10, n = 128$.

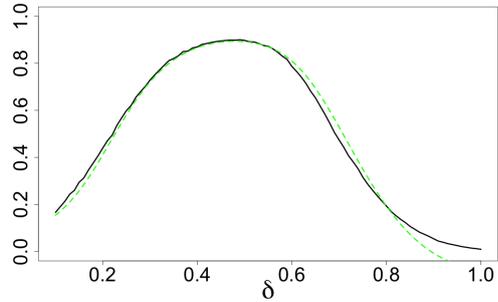


Figure 9.26: Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 10, n = 256$.

of points duplicate. This is reflected in Table 9.1 by comparing $n = 100$ and $n = 500$ for Design 2a; there is only a small reduction in r despite a large increase in n . Moreover, for values of $n \gg 2^d$, Design 2a provides a very inefficient covering.

- For $n \gg 2^d$, from looking at Design 1 with $\alpha = 0.5$ and $n = 500$, it would appear beneficial to choose $\alpha \in (0, 1)$ rather than $\alpha > 1$ or $\alpha = 0$.

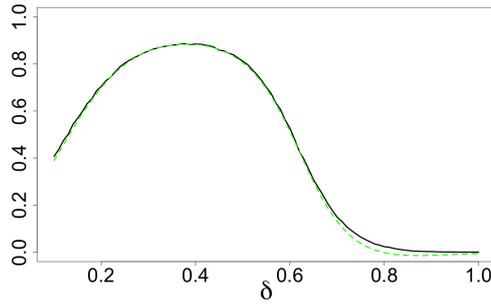


Figure 9.27: Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 20, n = 512$.

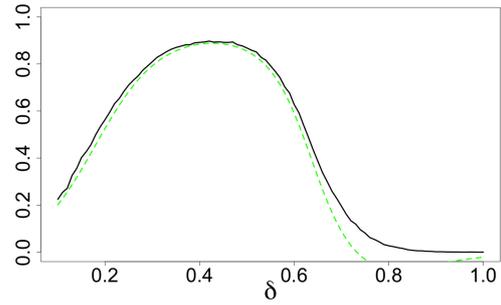


Figure 9.28: Design 2b: $C_d(\mathbb{Z}_n, r)$ and approximation (9.5.2); $d = 20, n = 2048$.

$d = 5$				
	$n = 25$	$n = 50$	$n = 100$	$n = 500$
Design 2a ($\alpha = 0$)	1.051 (0.44)	0.885 (0.50)	0.812 (0.50)	0.798 (0.50)
Design 1, $\alpha = 0.5$	1.072 (0.68)	0.905 (0.78)	0.770 (0.78)	0.540 (0.80)
Design 1, $\alpha = 1$	1.072 (0.78)	0.931 (0.86)	0.798 (0.98)	0.555 (1.00)
Design 1, $\alpha = 1.5$	1.091 (0.92)	0.950 (0.96)	0.820 (0.98)	0.589 (1.00)

Table 9.1: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 5$.

$d = 10$				
	$n = 500$	$n = 1000$	$n = 5000$	$n = 10000$
Design 2a ($\alpha = 0$)	1.228 (0.50)	1.135 (0.50)	1.073 (0.50)	1.071 (0.50)
Design 1, $\alpha = 0.5$	1.271 (0.69)	1.165 (0.73)	0.954 (0.76)	0.886 (0.78)
Design 1, $\alpha = 1$	1.297 (0.87)	1.194 (0.90)	0.992 (0.93)	0.917 (0.95)
Design 1, $\alpha = 1.5$	1.320 (1.00)	1.220 (1.00)	1.032 (1.00)	0.953 (1.00)

Table 9.2: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 10$.

Using approximations (9.3.10) and (9.4.5), in Figures 9.29–9.30 we depict $C_d(\mathbb{Z}_n, r)$ across δ for different choices of α . In Figures 9.29–9.30, the red line, green line, blue line and cyan line depict approximation (9.4.5) ($\alpha = 0$) and approximation (9.3.10) with $\alpha = 0.5, \alpha = 1$ and $\alpha = 1.5$ respectively. These figures demonstrate the clear benefit of choosing a smaller α , at least for these values of n and d .

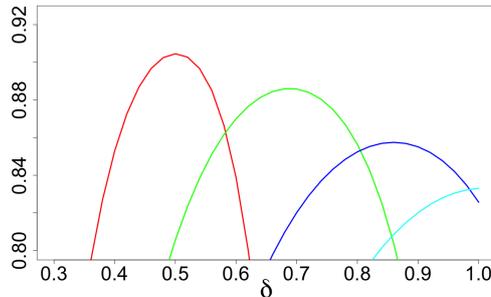


Figure 9.29: $d = 10, n = 512, r = 1.228$

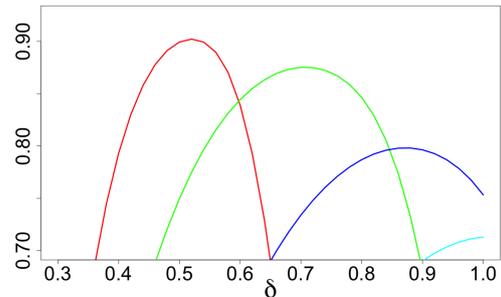


Figure 9.30: $d = 10, n = 1024, r = 1.13$

9.7 Quantization in a cube

9.7.1 Quantization error and its relation to weak covering

In this section, we will study the following characteristic of a design \mathbb{Z}_n .

Quantization error. Let $U = (u_1, \dots, u_d)$ be uniform random vector on $[-1, 1]^d$. The mean squared quantization error for a design $\mathbb{Z}_n = \{Z_1, \dots, Z_n\} \subset \mathbb{R}^d$ is defined by

$$\theta(\mathbb{Z}_n) = \mathbb{E}_U \varrho^2(U, \mathbb{Z}_n), \quad \text{where } \varrho(U, \mathbb{Z}_n) = \min_{Z_i \in \mathbb{Z}_n} \|U - Z_i\|. \quad (9.7.1)$$

If the design \mathbb{Z}_n is randomized then we consider the expected value $\mathbb{E}_{\mathbb{Z}_n} \theta(\mathbb{Z}_n)$ of $\theta(\mathbb{Z}_n)$ as the main characteristic without stressing this.

The mean squared quantization error $\theta(\mathbb{Z}_n)$ is related to our main quantity $C_d(\mathbb{Z}_n, r)$ defined in (9.1.1), in the following way:

$$\theta(\mathbb{Z}_n) = \int_{r \geq 0} r^2 dC_d(\mathbb{Z}_n, r). \quad (9.7.2)$$

This relation will allow us to use the approximations derived above for $C_d(\mathbb{Z}_n, r)$ in order to construct approximations for the quantization error $\theta(\mathbb{Z}_n)$.

9.7.2 Quantization error for Design 1

Using approximation (9.3.10) for the quantity $C_d(\mathbb{Z}_n, r)$, we obtain

$$\begin{aligned} \frac{d}{dr}(C_d(\mathbb{Z}_n, r)) &\cong f_{\alpha, \delta}(r) := \frac{n \cdot r}{\delta} \int_{-\infty}^{\infty} \varphi(s) \varphi(c_{s,r}) \psi_{2,\alpha}(s, r) \times \\ &\times \left[\frac{\sqrt{2\alpha+1}}{\sqrt{s'+k}} + c_{d,\alpha} \frac{\alpha \left(s' + \frac{d\delta^2(2\alpha-1)}{3(2\alpha+5)(2\alpha+1)} \right)}{(2\alpha+3)(s'+k)^2} \left\{ \delta(c_{s,r}^3 - c_{s,r}) - \frac{\sqrt{2\alpha+1}(r^2 - \frac{d\delta^2}{2\alpha+1} - s')}{\sqrt{s'+k}} \right\} \right] ds. \end{aligned} \quad (9.7.3)$$

By then using relation (9.7.2) we obtain the following approximation for the mean squared quantization error with Design 1:

$$\theta(\mathbb{Z}_n) \cong \int_0^{\infty} r^2 f_{\alpha, \delta}(r) dr. \quad (9.7.4)$$

By taking $\alpha = 1$ in (9.7.3) we obtain:

$$\begin{aligned} f_{1, \delta}(r) &:= \frac{n \cdot r}{\delta} \int_{-\infty}^{\infty} \varphi(s) \varphi(c_{s,r}) \psi_{2,1}(s, r) \times \\ &\times \left[\frac{\sqrt{3}}{\sqrt{s'+k}} + c_{d,1} \frac{\left(s' + \frac{d\delta^2}{63} \right)}{5(s'+k)^2} \left\{ \delta(c_{s,r}^3 - c_{s,r}) - \frac{\sqrt{3}(r^2 - \frac{d\delta^2}{3} - s')}{\sqrt{s'+k}} \right\} \right] ds. \end{aligned}$$

with $\psi_{2,1}$ defined in (9.3.11). The resulting approximation

$$\theta(\mathbb{Z}_n) \cong \int_0^{\infty} r^2 f_{1, \delta}(r) dr.$$

coincides with (7.5.2) of Chapter 7.

9.7.3 Quantization error for Design 2a

Using approximation (9.4.5) for the quantity $C_d(\mathbb{Z}_n, r)$, we have:

$$\frac{d}{dr}(C_d(\mathbb{Z}_n, r)) \cong f_{0,\delta;n}(r) := \frac{n \cdot r}{\delta} \int_{-\infty}^{\infty} \frac{\varphi(s)\varphi(c_{s,r})\psi_{4,n}(s, r)}{\sqrt{s'}} \left[1 + \left(1 + \frac{3}{d} \right) \frac{(2(s' - d/3)/\sqrt{5} + d/5)(6c_{s,r}^2 - c_{s,r}^4 - 3)}{12(s')^2} \right] ds, \quad (9.7.5)$$

where $\psi_{4,n}(\cdot)$ is defined in (9.4.6). From (9.7.2) we then obtain the following approximation for the mean squared quantization error with Design 2a:

$$\theta(\mathbb{Z}_n) \cong \int_0^{\infty} r^2 f_{0,\delta;n}(r) dr. \quad (9.7.6)$$

9.7.4 Quantization error for Design 2b

Similarly to (9.7.6), for Design 2b, we use the approximation

$$\theta(\mathbb{Z}_n) \cong \int_0^{\infty} r^2 f_{0,\delta;m}(r) dr. \quad (9.7.7)$$

where $f_{0,\delta,m}(r)$ is defined by (9.7.5) and $m = m_{n,d}$ is defined in (9.5.3).

9.7.5 Accuracy of approximations for quantization error and the δ -effect

In this section, we assess the accuracy of approximations (9.7.4), (9.7.6) and (9.7.7). Using a black line we depict $\mathbb{E}_{\mathbb{Z}_n}\theta(\mathbb{Z}_n)$ obtained via Monte Carlo simulations with 50,000 iterations. Depending on the value of α , in Figures 9.31–9.36 approximation (9.7.4) or (9.7.6) is shown using a red line. In Figures 9.41–9.44, approximation (9.7.7) is depicted with a red line. From the figures below we can see that all approximations are generally very accurate. Approximation (9.7.6) is much more accurate than approximation (9.7.4) across all choices of δ and n and this can be explained by the additional term taken in the general expansion; see Section 9.4.2. This high accuracy is also seen with approximation (9.7.7). The accuracy of approximation (9.7.4) seems to worsen for large δ , n and d not too large like $d = 20$, see Figures 9.33–9.34. For $d = 50$, all approximations are extremely accurate for all choices of δ and n . Figures 9.31–9.36 very clearly demonstrate the δ -effect implying that a sensible choice of δ is crucial for good quantization.

9.8 Comparative numerical studies of covering properties for several designs

Let us extend the range of designs considered above by adding the following two designs.

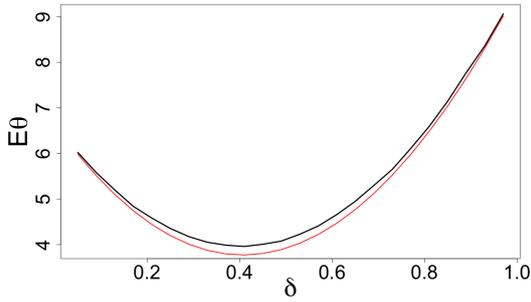


Figure 9.31: $\mathbb{E}\theta(Z_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 1$, $n = 500$

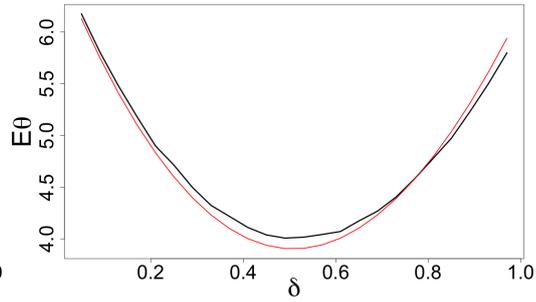


Figure 9.32: $\mathbb{E}\theta(Z_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 0.5$, $n = 500$

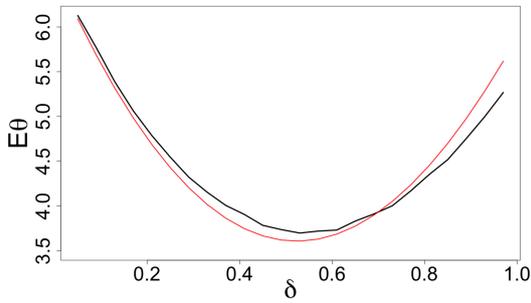


Figure 9.33: $\mathbb{E}\theta(Z_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 0.5$, $n = 1000$

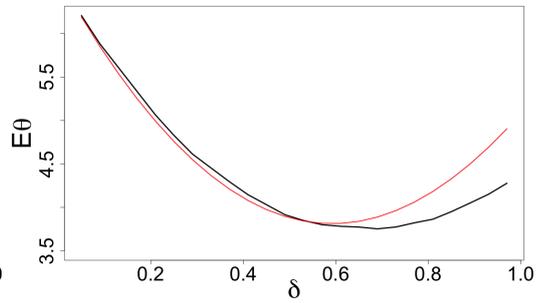


Figure 9.34: $\mathbb{E}\theta(Z_n)$ and approximation (9.7.4): $d = 20$, $\alpha = 1$, $n = 1000$

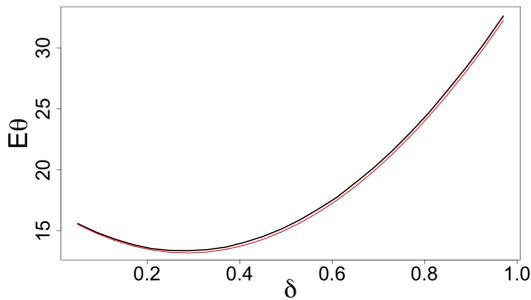


Figure 9.35: $\mathbb{E}\theta(Z_n)$ and approximation (9.7.4): $d = 50$, $\alpha = 0.1$, $n = 1000$

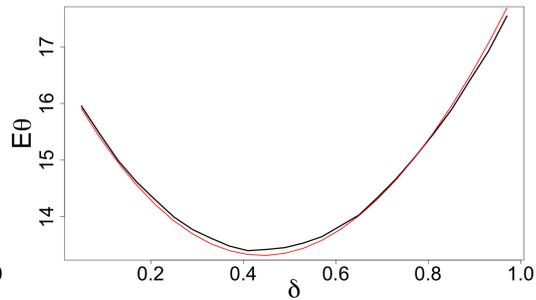


Figure 9.36: $\mathbb{E}\theta(Z_n)$ and approximation (9.7.4): $d = 50$, $\alpha = 1$, $n = 1000$

Design 3. Z_1, \dots, Z_n are taken from a low-discrepancy Sobol's sequence on the cube $[-\delta, \delta]^d$.

Design 4. Z_1, \dots, Z_n are taken from the minimum-aberration 2^{d-k} fractional factorial design on the vertices of the cube $[-\delta, \delta]^d$.

Unlike Designs 1, 2a, 2b and 3, Design 4 is non-adaptive and defined only for a

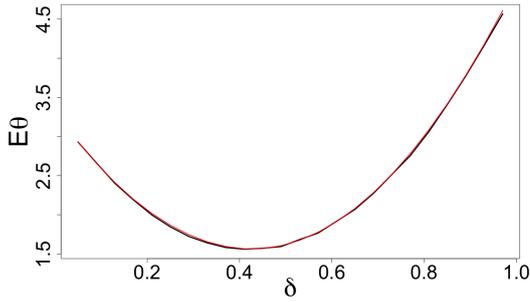


Figure 9.37: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 10, \alpha = 0, n = 100$

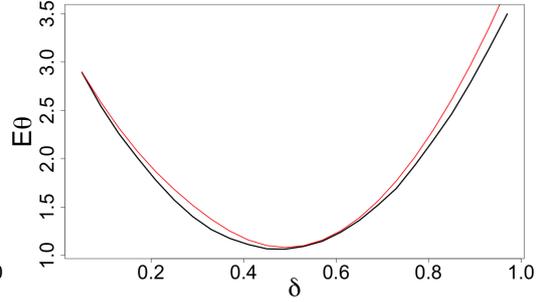


Figure 9.38: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 10, \alpha = 0, n = 500$

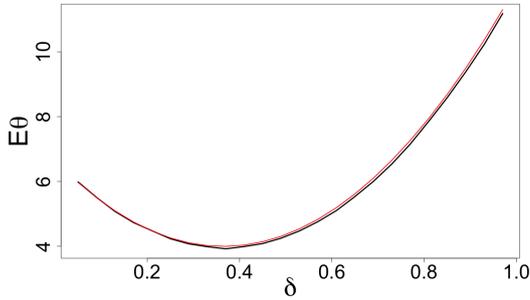


Figure 9.39: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 20, \alpha = 0, n = 500$

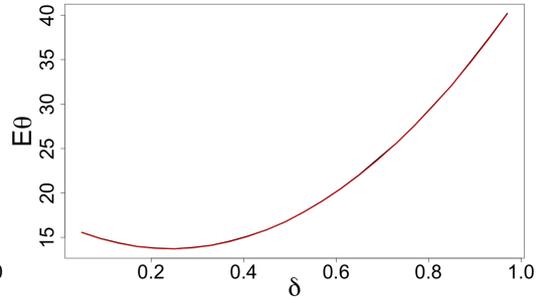


Figure 9.40: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.6): $d = 50, \alpha = 0, n = 500$

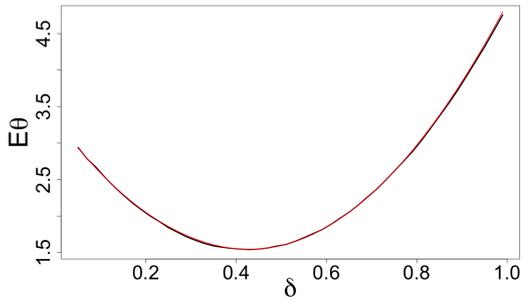


Figure 9.41: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 10, n = 100$

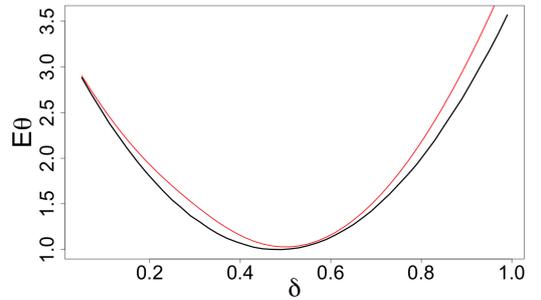


Figure 9.42: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 10, n = 500$

particular n of the form $n = 2^{d-k}$ with some $k \geq 0$. The author has included this design into the list of all designs as "the golden standard". In view of the numerical study in Chapter 7 and theoretical arguments in Chapter 8, Design 4 with $k = 1$ and optimal δ provides the best quantization the author could find; moreover, the author conjectures in Chapter 8 that Design 4 with $k = 1$ and optimal δ provides minimal normalized mean squared quantization error for all designs with $n \leq 2^d$. We repeat, Design 4 is defined for one particular value of n only.

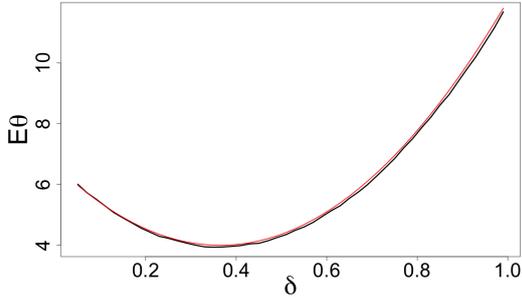


Figure 9.43: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 20, n = 500$

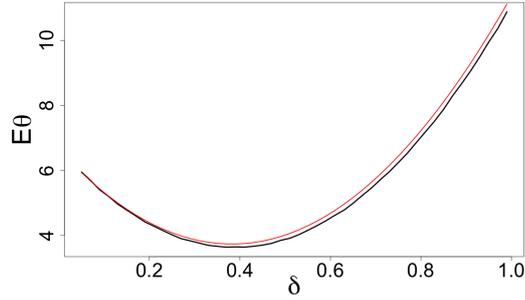


Figure 9.44: $\mathbb{E}\theta(\mathbb{Z}_n)$ and approximation (9.7.7): $d = 20, n = 1000$

9.8.1 Covering comparisons

In Tables 9.3–9.4, we present results of 50,000 Monte Carlo runs where we have computed the smallest values of r required to achieve the 0.9-coverage on average (on average, for Designs 1, 2a, 2b). The value inside the brackets shows the value of δ required to obtain the 0.9-coverage.

$d = 10$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Design 1, $\alpha = 0.5$	1.629 (0.58)	1.505 (0.65)	1.270 (0.72)	1.165 (0.75)
Design 1, $\alpha = 1.5$	1.635 (0.80)	1.525 (0.88)	1.310 (1.00)	1.210 (1.00)
Design 2a	1.610 (0.38)	1.490 (0.46)	1.228 (0.50)	1.132 (0.50)
Design 2b	1.609 (0.41)	1.475 (0.43)	1.178 (0.49)	1.075 (0.50)
Design 3	1.595 (0.72)	1.485 (0.80)	1.280 (0.85)	1.170 (0.88)
Design 3, $\delta = 1$	1.678 (1.00)	1.534 (1.00)	1.305 (1.00)	1.187 (1.00)
Design 4	1.530 (0.44)	1.395 (0.48)	1.115 (0.50)	1.075 (0.50)

Table 9.3: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 10$.

$d = 20$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Design 1, $\alpha = 0.5$	2.540 (0.44)	2.455 (0.48)	2.285 (0.55)	2.220 (0.60)
Design 1, $\alpha = 1.5$	2.545 (0.60)	2.460 (0.65)	2.290 (0.76)	2.215 (0.84)
Design 2a	2.538 (0.28)	2.445 (0.30)	2.270 (0.36)	2.180 (0.42)
Design 2b	2.538 (0.29)	2.445 (0.30)	2.253 (0.37)	2.173 (0.42)
Design 3	2.520 (0.50)	2.445 (0.60)	2.285 (0.68)	2.196 (0.72)
Design 3, $\delta = 1$	2.750 (1.00)	2.656 (1.00)	2.435 (1.00)	2.325 (1.00)
Design 4	2.490 (0.32)	2.410 (0.35)	2.220 (0.40)	2.125 (0.44)

Table 9.4: Values of r and δ (in brackets) to achieve 0.9 coverage for $d = 20$.

From Tables 9.3–9.4 we can draw the following conclusions:

- Designs 2a and especially 2b provide very high quality coverage (on average) whilst being online procedures (that is, nested designs);
- Design 2b has significant benefits over Design 2a for values of n close to 2^d ;

- properly δ -tuned deterministic non-nested Design 4 provides superior covering;
- coverage properties of δ -tuned low-discrepancy sequences are much better than of the original low-discrepancy sequences;
- coverage of an unadjusted low-discrepancy sequence is poor.

In Figures 9.45–9.46, after fixing n and δ , we plot $C_d(\mathbb{Z}_n, r)$ as a function of r for the following designs: Design 1 with $\alpha = 1$ (red line), Design 2a (blue line), Design 2b (green line) and Design 3 with $\delta = 1$ (black line). For Design 1 with $\alpha = 1$, Design 2a and Design 2b, we have used approximations (9.3.11), (9.4.6) and (9.5.2) respectively to depict $C_d(\mathbb{Z}_n, r)$ whereas for Design 3, we have used Monte Carlo simulations with 50,000 iterations. For the first three designs, depending of the choice of n , the value of δ has been fixed based on the optimal value for quantization; these are the values inside the brackets in Tables 9.5–9.6.

From Figure 9.45, we see that Design 2b is superior and uniformly dominates all other designs for this choice of d and n (at least when the level of coverage is greater than $1/2$). In Figure 9.46, since $n \ll 2^d$, the values of $C_d(\mathbb{Z}_n, r)$ for Designs 2a and 2b practically coincide and the green line hides under the blue. In both figures we see that Design 3 with an unadjusted δ provides a very inefficient covering.

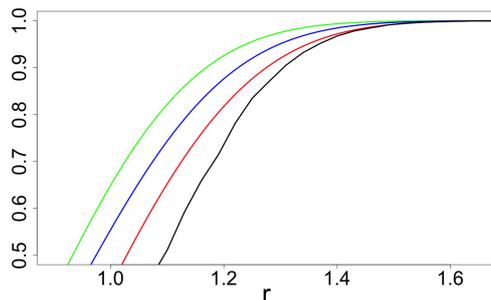


Figure 9.45: $C_d(\mathbb{Z}_n, r)$ as a function of r for several designs: $d = 10$, $n = 512$

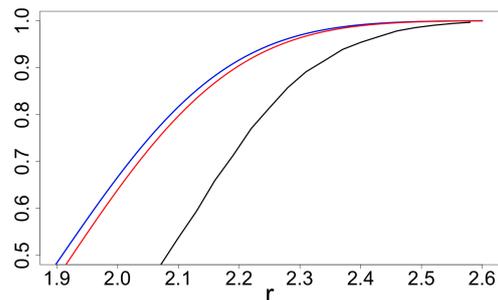


Figure 9.46: $C_d(\mathbb{Z}_n, r)$ as a function of r for several designs: $d = 20$, $n = 1024$

9.8.2 Quantization comparisons

As follows from results of [79, Ch.6], for efficient covering schemes the order of convergence of the covering radius to 0 as $n \rightarrow \infty$ is $n^{-1/d}$. Therefore, for the mean squared distance (which is the quantization error) we should expect the order $n^{-2/d}$ as $n \rightarrow \infty$. Therefore, for sake of comparison of quantization errors θ_n across n we renormalize this error from $\mathbb{E}\theta_n$ to $n^{2/d}\mathbb{E}\theta_n$.

In Figure 9.47–9.6, we present the minimum value of $n^{2/d}\mathbb{E}\theta_n$ for a selection of designs. In these tables, the value within the brackets corresponds to the value of δ where the minimum of $n^{2/d}\mathbb{E}\theta_n$ was obtained.

In Figure 9.47, we depict the c.d.f.'s for the distance $\varrho(X, \mathbb{Z}_n)$ for Design 2a with $\delta = 0.5$ (in red) and Design 3 with $\delta = 0.8$ (in black). We can see that for $d = 10$ and $n = 512$, Design 2a stochastically dominates Design 3. The style of Figure 9.48

$d = 10$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Design 1, $\alpha = 0.5$	4.072 (0.56)	4.013 (0.60)	3.839 (0.68)	3.770 (0.69)
Design 1, $\alpha = 1$	4.153 (0.68)	4.105 (0.72)	3.992 (0.80)	3.925 (0.84)
Design 1, $\alpha = 1.5$	4.164 (0.80)	4.137 (0.86)	4.069 (0.96)	4.026 (0.98)
Design 2a	3.971 (0.38)	3.866 (0.44)	3.670 (0.48)	3.704 (0.50)
Design 2b	3.955 (0.40)	3.798 (0.44)	3.453 (0.48)	3.348 (0.50)
Design 3	3.998 (0.68)	3.973 (0.76)	3.936 (0.80)	3.834 (0.82)
Design 3, $\delta = 1$	4.569 (1.00)	4.425 (1.00)	4.239 (1.00)	4.094 (1.00)
Design 4	3.663 (0.40)	3.548 (0.44)	3.219 (0.49)	3.348 (0.50)

Table 9.5: Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across selected designs; $d = 10$.

$d = 20$				
	$n = 64$	$n = 128$	$n = 512$	$n = 1024$
Design 1, $\alpha = 0.5$	7.541 (0.40)	7.515 (0.44)	7.457 (0.52)	7.421 (0.54)
Design 1, $\alpha = 1$	7.552 (0.52)	7.563 (0.56)	7.528 (0.64)	7.484 (0.68)
Design 1, $\alpha = 1.5$	7.561 (0.60)	7.571 (0.64)	7.556 (0.74)	7.527 (0.78)
Design 2a	7.488 (0.30)	7.461 (0.33)	7.346 (0.35)	7.248 (0.39)
Design 2b	7.487 (0.29)	7.458 (0.34)	7.345 (0.36)	7.234 (0.40)
Design 3	7.445 (0.48)	7.464 (0.56)	7.487 (0.64)	7.453 (0.66)
Design 3, $\delta = 1$	9.089 (1.00)	9.133 (1.00)	8.871 (1.00)	8.681 (1.00)
Design 4	7.298 (0.32)	7.270 (0.33)	7.133 (0.36)	7.016 (0.40)

Table 9.6: Minimum value of $n^{2/d}\mathbb{E}\theta_n$ and δ (in brackets) across selected designs; $d = 20$.

is the same as figure Figure 9.47, however we set $n = 1024$ and Design 2a is replaced with Design 2b with $\delta = 0.5$ (we also set $\delta = 0.82$ for Design 3). Here we see a very clear stochastic dominance of the Design 2b over Design 4. All findings are consistent with Tables 9.5 and 9.6. In Figures 9.47 and 9.48, values of the parameter δ for all designs are chosen as numerically optimal, in accordance with Table 9.5.

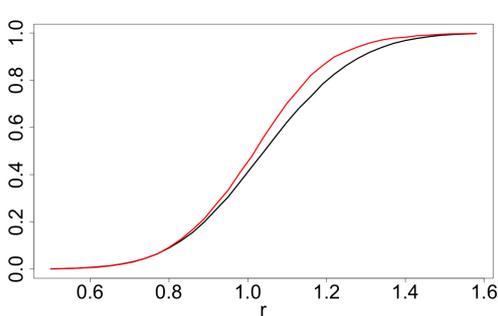


Figure 9.47: $d = 10, n = 512$: Design 2a with $\delta = 0.5$ stochastically dominates Design 3 with $\delta = 0.8$.

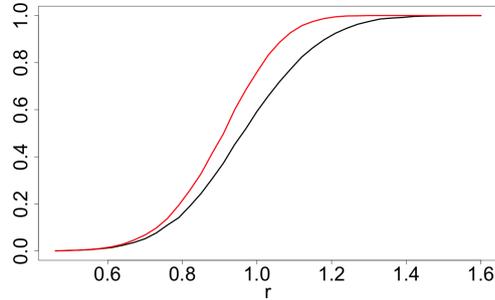


Figure 9.48: $d = 10, n = 1024$: Design 2b with $\delta = 0.5$ stochastically dominates Design 3 with $\delta = 0.82$.

We make the following conclusions from analyzing results of this section:

- Designs 2a and 2b provide very good quantization per point. As expected, Design 2b is superior over Design 2a when n is close to 2^d ; see Table 9.5.
- Properly δ -tuned non-nested Design 4 provides the best quantization per point of all designs considered.
- Properly δ -tuned Design 3 is comparable in performance to Design 1 but it is not as efficient as Designs 2a, 2b and 4.

Note that for the special case of $n = 2^{d-1}$, for Design 4 we could use the results of Chapter 8 to obtain the exact value of $n^{2/d}\mathbb{E}\theta_n$. In particular, using Theorem 8.2.2 with the optimal value of δ obtained from Corollary 8.2.1, we obtain the minimal value of $n^{2/d}\mathbb{E}\theta_n$ is

$$n^{2/d}\mathbb{E}\theta_n = 2^{2-2/d} \left[\frac{d}{12} + \frac{1}{d+1} \right].$$

This is consistent with the $n = 512$ column of Table 9.5. The figures in this column do not violate the conjecture stated in Chapter 8 for Design 4 with $n = 2^{d-1}$.

9.9 Covering and quantization in the d -simplex

9.9.1 Characteristics of interest

Consider the standard orthogonal d -simplex

$$\mathcal{S}_d := \left\{ (u_1, u_2, \dots, u_d) \in \mathbb{R}^d \mid \sum_{i=1}^d u_i \leq 1 \text{ and } u_i \geq 0 \text{ for all } i \right\}$$

with $\text{vol}(\mathcal{S}_d) = 1/d!$. For a design $\mathbb{Z}_n = \{Z_1, \dots, Z_n\}$, consider the following two characteristics:

- (a) the proportion of the simplex \mathcal{S}_d covered by $\mathcal{B}_d(\mathbb{Z}_n, r)$:

$$C_d(\mathbb{Z}_n, r) := d! \text{vol}(\mathcal{S}_d \cap \mathcal{B}_d(\mathbb{Z}_n, r)), \quad (9.9.1)$$

- (b) $\theta(\mathbb{Z}_n) = \mathbb{E}_U \min_{i=1, \dots, n} \|U - Z_i\|^2$, the mean squared quantization error for \mathbb{Z}_n , where $U = (u_1, \dots, u_d)$ is a random vector uniformly distributed in \mathcal{S}_d .

In this section, we investigate whether the δ -effect seen in Sections 9.6, 9.7.5 and 9.8 for the cube is present for the simplex \mathcal{S}_d . We will consider two possible ways of scaling points in \mathcal{S}_d . Define the two δ -simplices $\mathcal{S}_{d,1}^{(\delta)}$ and $\mathcal{S}_{d,2}^{(\delta)}$ as follows:

$$\mathcal{S}_{d,1}^{(\delta)} := \delta \cdot \mathcal{S}_d,$$

$$\mathcal{S}_{d,2}^{(\delta)} := \left\{ (u_1, u_2, \dots, u_d) \in \mathbb{R}^d \mid \sum_{i=1}^d u_i \leq \frac{d+\delta}{d+1} \text{ and } u_i \geq \frac{1-\delta}{d+1} \text{ for all } i \right\}.$$

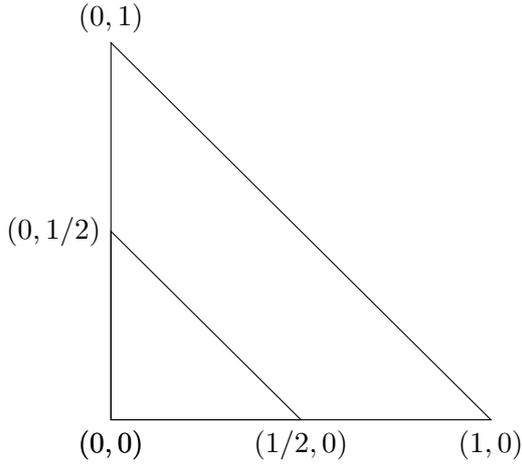


Figure 9.49: \mathcal{S}_d and $\mathcal{S}_{d,1}^{(\delta)}$ with $d = 2$ and $\delta = 0.5$

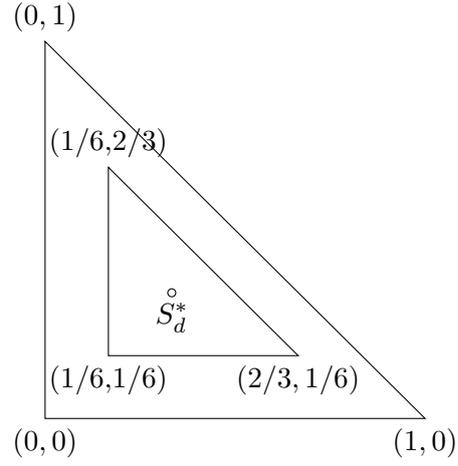


Figure 9.50: \mathcal{S}_d and $\mathcal{S}_{d,2}^{(\delta)}$ with $d = 2$ and $\delta = 0.5$

By construction, the value of δ in $\mathcal{S}_{d,2}^{(\delta)}$ scales the simplex around its centroid $S_d^* = \left(\frac{1}{d+1}, \frac{1}{d+1}, \dots, \frac{1}{d+1}\right)$, where for $\delta = 1$, we have $\mathcal{S}_{d,2}^{(\delta)} = \mathcal{S}_d$. Simple depictions of $\mathcal{S}_{d,1}^{(\delta)}$ and $\mathcal{S}_{d,2}^{(\delta)}$ are given in Figures 9.49–9.50.

We will numerically assess covering and quantization characteristics for the following two designs.

Design S1. Z_1, \dots, Z_n are i.i.d. random vectors uniformly distributed in the δ -scaled simplex $\mathcal{S}_{d,1}^{(\delta)}$, where $\delta \in [0, 1]$ is a parameter.

Design S2. Z_1, \dots, Z_n are i.i.d. random vectors uniformly distributed in the δ -scaled simplex $\mathcal{S}_{d,2}^{(\delta)}$, where $\delta \in [0, 1]$ is a parameter.

To simulate points Y uniformly distributed in the simplex \mathcal{S}_d , one can simply generate d i.i.d. uniformly distributed points in $[0, 1]$, add 0 and 1 to the collection of points and take the first d spacings (out of the total number $d + 1$ of these spacings). Points $Y' = \delta Y$ and $Y'' = \delta \cdot (Y - S_d^*) + S_d^*$ are then uniform in $\mathcal{S}_{d,1}^{(\delta)}$ and $\mathcal{S}_{d,2}^{(\delta)}$ respectively. This procedure can be easily performed in R using the package ‘uniformly’.

9.9.2 Numerical investigation of the δ -effect for d -simplex

Using the above procedure, we numerically study characteristics of Designs S1 and S2. In Figures 9.51–9.54 we plot $C_d(\mathbb{Z}_n, r)$ as a functions of $\delta \in [0, 1]$ across n, r and d for Design S1. The corresponding results for Design S2 are given in Figures 9.55–9.58. In Figures 9.59–9.60 and Figures 9.61–9.62, we depict $\mathbb{E}\theta(\mathbb{Z}_n)$ for Designs S1 and S2 respectively for different n and d . In each figure we plot values of $\mathbb{E}\theta(\mathbb{Z}_n)$ for different values of r ; a step in r increase gives the next curve up.

From the above figures, we arrive at the following conclusions:

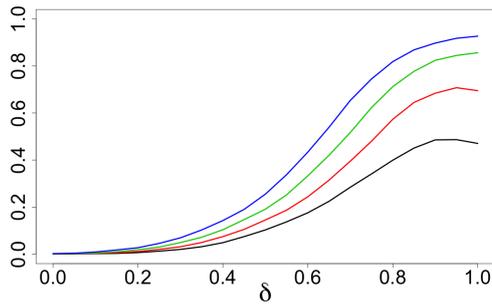


Figure 9.51: $C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 5$, $n = 128$, r from 0.11 to 0.17 increasing by 0.02.

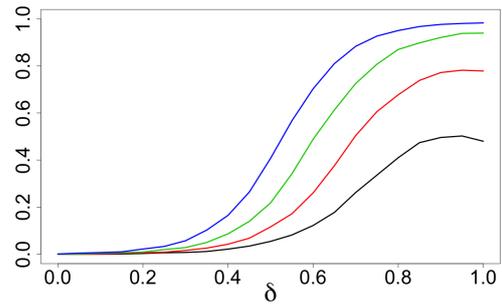


Figure 9.52: $C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 10$, $n = 512$, r from 0.13 to 0.19 increasing by 0.02.

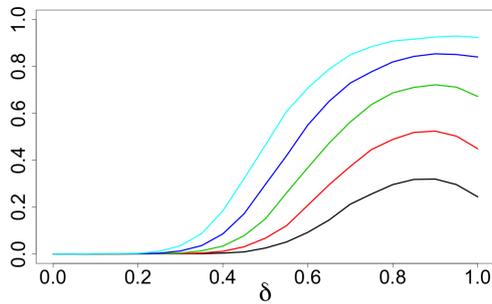


Figure 9.53: $C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 20$, $n = 1024$, r from 0.13 to 0.17 increasing by 0.01.

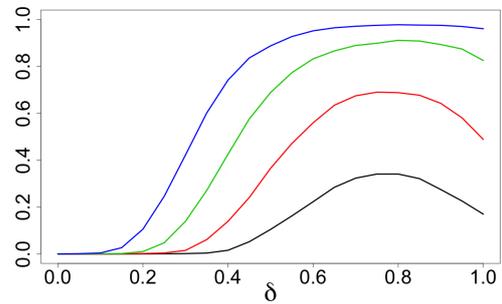


Figure 9.54: $C_d(\mathbb{Z}_n, r)$ for Design S1: $d = 50$, $n = 1024$, r from 0.12 to 0.15 increasing by 0.01.

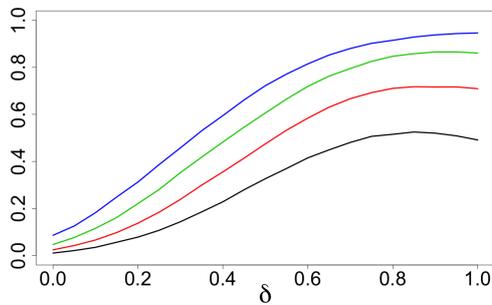


Figure 9.55: $C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 5$, $n = 128$, r from 0.11 to 0.17 increasing by 0.02.

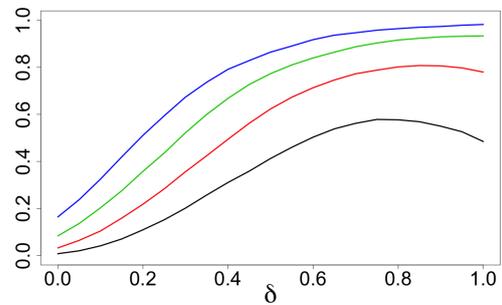


Figure 9.56: $C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 10$, $n = 512$, r from 0.13 to 0.19 increasing by 0.02.

- The δ -effect for the simplex is much less prominent than for the cube.
- Between Designs S1 and S2, the δ -effect is more apparent for Design S2; for example, compare Figure 9.60 with Figure 9.62.

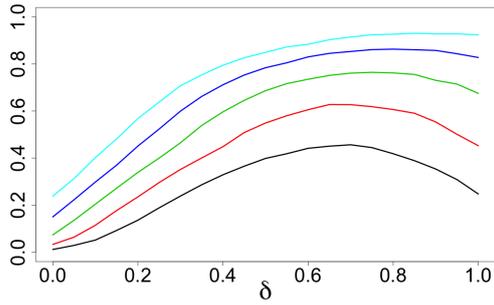


Figure 9.57: $C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 20, n = 1024, r$ from 0.13 to 0.17 increasing by 0.01.

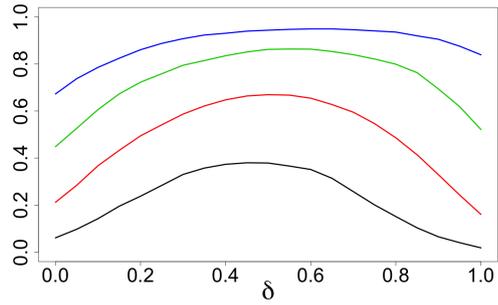


Figure 9.58: $C_d(\mathbb{Z}_n, r)$ for Design S2: $d = 50, n = 1024, r$ from 0.11 to 0.14 increasing by 0.01.

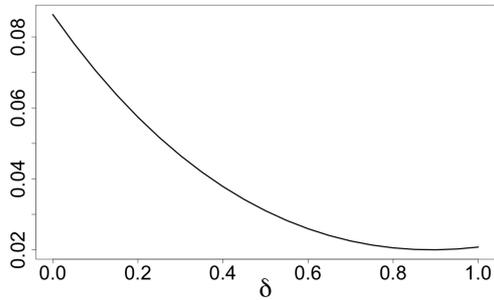


Figure 9.59: $\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S1: $d = 20, n = 1024$.

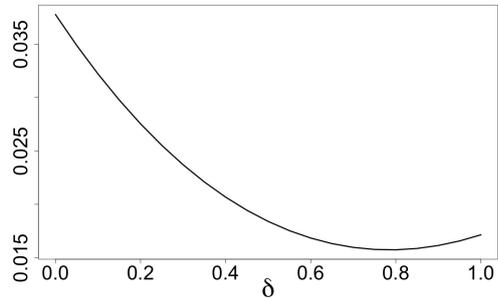


Figure 9.60: $\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S1: $d = 50, n = 1024$.

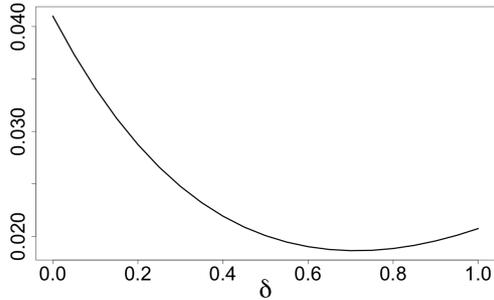


Figure 9.61: $\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S2: $d = 20, n = 1024$.

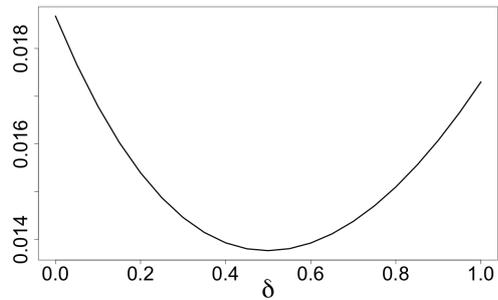


Figure 9.62: $\mathbb{E}\theta(\mathbb{Z}_n)$ for Design S2: $d = 50, n = 1024$.

9.10 Appendix: An auxiliary lemma

Lemma 9.10.1 Let $\delta > 0, u \in \mathbb{R}$ and $\eta_{u,\delta}$ be a r.v. $\eta_{u,\delta} = (\xi - u)^2$, where r.v. $\xi \in [-\delta, \delta]$ has $Beta_\delta(\alpha, \alpha)$ distribution with density

$$p_{\alpha,\delta}(t) = \frac{(2\delta)^{1-2\alpha}}{Beta(\alpha, \alpha)} [\delta^2 - t^2]^{\alpha-1}, \quad -\delta < t < \delta, \alpha > 0; \quad (9.10.1)$$

$Beta(\cdot, \cdot)$ is the Beta-function. The r.v. $\eta_{u,\delta}$ is concentrated on the interval $[(\max(0, \delta - |u|))^2, (\delta + |u|)^2]$. Its first three central moments are:

$$\begin{aligned}\mu_{u,\delta}^{(1)} &= E\eta_{u,\delta} = u^2 + \frac{\delta^2}{2\alpha + 1}, \\ \mu_{u,\delta}^{(2)} &= \text{var}(\eta_{u,\delta}) = \frac{4\delta^2}{2\alpha + 1} \left[u^2 + \frac{\delta^2\alpha}{(2\alpha + 1)(2\alpha + 3)} \right], \\ \mu_{u,\delta}^{(3)} &= E[\eta_{u,\delta} - E\eta_{u,\delta}]^3 = \frac{48\alpha\delta^4}{(2\alpha + 1)^2(2\alpha + 3)} \left[u^2 + \frac{\delta^2(2\alpha - 1)}{3(2\alpha + 5)(2\alpha + 1)} \right].\end{aligned}$$

In the limiting case $\alpha = 0$, where the r.v. ξ is concentrated at two points $\pm\delta$ with equal weights, we obtain: $\mu_{u,\delta}^{(1)} = E\eta_{u,\delta} = u^2 + \delta^2$ and

$$\mu_{u,\delta}^{(2k)} = [2\delta u]^{2k}, \quad \mu_{u,\delta}^{(2k+1)} = 0, \quad \text{for } k = 1, 2, \dots \quad (9.10.2)$$

Chapter 10

Summary of this thesis

10.0.1 Summary of Part one and next steps

In Part one of this thesis, we considered the topic of change-point detection. We addressed the problem of detecting a transient change in distributions of i.i.d. random variables. The majority of research in Part one focused on deriving important change-point detection quantities, like average run length and power, for the MOSUM change-point procedure. In Chapter 1, we provided an introduction to the change-point problem considered in this thesis and offered a survey of the current state of the field. It also provided a survey of the main findings of the author and can be used as a summary of Part one of this thesis. Here it was highlighted that a lot of research has been focused on detecting transient changes in i.i.d. random variables with and without nuisance parameters in the offline change point setting, where observations do not arrive sequentially. However, to the author's knowledge there is no immediate way to implement these impressive offline results for the online change-point problem. In Chapter 2, we derived accurate approximations for boundary crossing probabilities for the MOSUM statistic and derived a very accurate approximation for average run length. The key methodology used in this thesis is to use boundary crossing probabilities for the MOSUM test in continuous time which can be computed explicitly, and subsequently correct the results for discrete time by generalising the sequential analysis results of David Siegmund. In Chapter 3, we studied constants related to continuous time MOSUM procedure. In Chapter 4, we derived previously unseen boundary-crossing probabilities for the continuous time MOSUM statistic. By deriving results for the first-passage probability when the barrier under consideration is piece-wise linear, we could evaluate the power of the change-point procedure in continuous time. Subsequently in Chapter 5, we corrected the expressions for power for discreteness to obtain accurate approximations to power in discrete time. In Chapter 6, we discussed boundary-crossing probabilities related to the popular Singular Spectrum Analysis change-point detection procedure and derived a number of approximations.

For future research, different forms of transient changes will be considered. For certain problems, the so-called epidemic change in distributions considered in this thesis (constant change in mean) can sometimes be deemed impractical. Instead, detecting a linear change in mean is potentially more useful. This will have immediate

implications to the SSA change-point detection algorithm mentioned in Chapter 6. The algorithm can be easily adapted to detect transient changes in a faster manner by considering a certain weighted distance. By likelihood ratio arguments, one can show that weights should approximately correspond to the change you expect, at least when considering transient changes in mean.

10.0.2 Summary of Part two and next steps

In Part two, we considered the problem of covering and quantization of high dimensional sets. In particular, we considered the covering and quantization of a d -dimensional cube by n balls with reasonably large d (10 or more) and reasonably small n . We focused on the notion of weak covering (say 99% covering of the cube) and the results of Part two establish that efficient covering schemes have several important properties which are not seen in small dimensions and in asymptotical considerations, for very large n . One of these properties can be termed ‘do not try to cover the vertices’ as the vertices of the cube and their close neighbourhoods are very hard to cover and for large d there are too many of them. The results of Chapter 8 highlight this, where we proved that in high dimensions for an arrangement of points related to a famous lattice, we could cover 99.99% of the cube with a radius $1/\sqrt{3}$ times smaller than the one required for full coverage.

For future research, the performance of other lattices constrained to the cube should be theoretically studied. This will be important in checking the validity of the conjecture stated in Chapter 8. Instead of covering the d -dimensional cube, research into covering of discrete sets in high dimensions should also be explored. Discrete sets like the vertices of the d -dimensional cube have immediate connections to the field of group testing; it would be rewarding if the notion of weak covering introduced in this thesis can be adapted to enhance results in this field.

Bibliography

- [1] R. Adler and J. Taylor. *Random Fields and Geometry*. Springer, 2007.
- [2] D. Aldous. *Probability Approximations via the Poisson Clumping Heuristic*. Springer Science & Business Media, 1989.
- [3] B. Bakhache and I. Nikiforov. Reliable detection of faults in measurement systems. *International Journal of adaptive control and signal processing*, 14(7):683–700, 2000.
- [4] M. Basseville and I. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall Inc, 1993.
- [5] P. Bauer and P. Hackl. The use of MOSUMS for quality control. *Technometrics*, 20(4):431–436, 1978.
- [6] P. Bauer and P. Hackl. An extension of the MOSUM technique for quality control. *Technometrics*, 22(1):1–7, 1980.
- [7] J. Beekman. Asymptotic distributions for the Ornstein-Uhlenbeck process. *Journal of Applied Probability*, 12(1):107–114, 1975.
- [8] M. Bell. Information theory and radar waveform design. *IEEE Transactions on Information Theory*, 39(5):1578–1597, 1993.
- [9] A. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49:122–136, 1941.
- [10] A. Bianchi, L. Mainardi, E. Petrucci, M. Signorini, M. Mainardi, and S. Cerutti. Time-variant power spectrum analysis for the detection of transient episodes in HRV signal. *IEEE Transactions on Biomedical Engineering*, 40(2):136–144, 1993.
- [11] W. Bischoff and A. Gegg. Boundary crossing probabilities for (q, d) -Slepian-processes. *Statistics & Probability Letters*, 118:139–144, 2016.
- [12] A. Blum, J. Hopcroft, and R. Kannan. *Foundations of data science*. Cambridge University Press, 2020.
- [13] W. Böhm and S. Mohanty. On the Karlin-Mcgregor theorem and applications. *The Annals of Applied Probability*, 7(2):314–325, 1997.

- [14] G. Box and S. Hunter. The 2^{k-p} fractional factorial designs. *Technometrics*, 3 (3):311–351, 1961.
- [15] E. Brodsky and B. Darkhovsky. *Nonparametric methods in change point problems*. Springer Science & Business Media, 2013.
- [16] Y. Chen, T. Wang, and R. Samworth. High-dimensional, multiscale online changepoint detection. *arXiv preprint arXiv:2003.03668*, 2020.
- [17] H. Cho and P. Fryzlewicz. Multiple-change-point detection for high dimensional time series via sparsified binary segmentation. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 77(2):475–507, 2015.
- [18] C. Chu, K. Hornik, and C. Kaun. MOSUM tests for parameter constancy. *Biometrika*, 82(3):603–617, 1995.
- [19] J. Conway and N. Sloane. *Sphere packings, lattices and groups*. Springer Science & Business Media, 2013.
- [20] H. Cramer. A limit theorem for the maximum values of certain stochastic processes. *Theory of Probability & Its Applications*, 10(1):126–128, 1965.
- [21] P. Deng. Boundary non-crossing probabilities for Slepian process. *Statistics & Probability Letters*, 122:28–35, 2017.
- [22] J. Durbin. The first-passage density of a continuous Gaussian process to a general boundary. *Journal of Applied Probability*, 22(1):99–122, 1985.
- [23] B. Eichinger and C. Kirch. A MOSUM procedure for the estimation of multiple random change points. *Bernoulli*, 24(1):526–564, 2018.
- [24] X. Fang, J. Li, and D. Siegmund. Segmentation and estimation of change-point models: false positive control and confidence regions. *The Annals of Statistics*, 48(3):1615–1647, 2020.
- [25] P. Fryzlewicz. Wild binary segmentation for multiple change-point detection. *Annals of Statistics*, 42(6):2243–2281, 2014.
- [26] P. Fryzlewicz. Detecting possibly frequent change-points: Wild Binary Segmentation 2 and steepest-drop model selection. *Journal of the Korean Statistical Society*, 49(4):1027–1070, 2020.
- [27] W. Fulton and J. Harris. *Representation theory: a first course*. Springer Science & Business Media, 2013.
- [28] A. Genz and F. Bretz. *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics. Springer-Verlag, Heidelberg, 2009. ISBN 978-3-642-01688-2.

- [29] A. Genz, F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, and T. Hothorn. *mvtnorm: Multivariate Normal and t Distributions*. 2018. URL <https://CRAN.R-project.org/package=mvtnorm>. R package version 1.0-8: ‘<https://CRAN.R-project.org/package=mvtnorm>’.
- [30] J. Glaz and B. Johnson. Boundary crossing for moving sums. *Journal of Applied Probability*, 25(1):81–88, 1988.
- [31] J. Glaz and J. Naus. Tight bounds and approximations for scan statistic probabilities for discrete data. *The Annals of Applied Probability*, 1(2):306–318, 1991.
- [32] J. Glaz, J. Naus, and S. Wallenstein. *Scan statistics*. Springer, 2001.
- [33] J. Glaz, V. Pozdnyakov, and S. Wallenstein. *Scan statistics: Methods and applications*. Springer Science & Business Media, 2009.
- [34] J. Glaz, J. Naus, and X. Wang. Approximations and inequalities for moving sums. *Methodology and Computing in Applied Probability*, 14(3):597–616, 2012.
- [35] N. Golyandina and A. Zhigljavsky. *Singular Spectrum Analysis for Time Series*. Springer Briefs in Statistics. Springer, 2013.
- [36] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky. *Analysis of Time Series Structure: SSA and related techniques*, volume 90 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC, London, UK, 2001.
- [37] L. Gordon and M. Pollak. An efficient sequential nonparametric scheme for detecting a change of distribution. *The Annals of Statistics*, 22(2):763–804, 1994.
- [38] U. Grenander. *Abstract inference*. John Wiley & Sons, 1981.
- [39] B. Guépié, L. Fillatre, and I. Nikiforov. Sequential detection of transient changes. *Sequential Analysis*, 31(4):528–547, 2012.
- [40] B. Guépié, L. Fillatre, and I. Nikiforov. Detecting a suddenly arriving dynamic profile of finite duration. *IEEE Transactions on Information Theory*, 63(5):3039–3052, 2017.
- [41] C. Han, P. Willett, and D. Abraham. Some methods to evaluate the performance of Page’s test as used to detect transient signals. *IEEE transactions on signal processing*, 47(8):2112–2127, 1999.
- [42] A. Harper. Pickands’ constant H_α does not equal $1/\Gamma(1/\alpha)$, for small α . *Bernoulli*, 23(1):582–602, 2017.
- [43] M. Hogan and D. Siegmund. Large deviations for the maxima of some random fields. *Advances in Applied Mathematics*, 7(1):2–22, 1986.
- [44] S. Janson. Random coverings in several dimensions. *Acta Mathematica*, 156:83–118, 1986.

- [45] J. Januszewski and M. Lassak. On-line covering the unit cube by cubes. *Discrete & Computational Geometry*, 12(4):433–438, 1994.
- [46] S. Joe and F. Kuo. Constructing Sobol sequences with better two-dimensional projections. *SIAM Journal on Scientific Computing*, 30(5):2635–2654, 2008.
- [47] M. Johnson, L. Moore, and D. Ylvisaker. Minimax and maximin distance designs. *Journal of statistical planning and inference*, 26(2):131–148, 1990.
- [48] Z. Kabluchko. Extreme-value analysis of standardized Gaussian increments. *arXiv preprint arXiv:0706.1849*, 2007.
- [49] S. Karlin and J. McGregor. Coincidence probabilities. *Pacific Journal of Mathematics*, 9(4):1141–1164, 1959.
- [50] M. Katori. O’Connell’s process as a vicious Brownian motion. *Physical Review E*, 84(6):061144, 2011.
- [51] M. Katori. Reciprocal time relation of noncolliding Brownian motion with drift. *Journal of Statistical Physics*, 148(1):38–52, 2012.
- [52] M. Katori and H. Tanemura. Scaling limit of vicious walks and two-matrix model. *Physical Review E*, 66(1):011105, 2002.
- [53] M. Katori and H. Tanemura. Noncolliding Brownian motion and determinantal processes. *Journal of Statistical Physics*, 129(5-6):1233–1277, 2007.
- [54] M. Katori and H. Tanemura. Noncolliding processes, matrix-valued processes and determinantal processes. *arXiv preprint arXiv:1005.0533*, 2010.
- [55] M. Katori, T. Nagao, and H. Tanemura. Infinite systems of non-colliding Brownian particles. In *Stochastic analysis on large scale interacting systems*, volume 39, pages 283–306. Mathematical Society of Japan, 2004.
- [56] K. Korkas and P. Fryzlewicz. Multiple change-point detection for non-stationary time series using wild binary segmentation. *Statistica Sinica*, 27(1):287–311, 2017.
- [57] F. Kuo, S. Joe, M. Matsumoto, S. Mori, and M. Saito. Sobol Sequences with Better Two-Dimensional Projections. <https://CRAN.R-project.org/package=SobolSequence>, 2017.
- [58] W. Kuperberg. On-line covering a cube by a sequence of cubes. *Discrete & Computational Geometry*, 12(1):83–90, 1994.
- [59] T. Lai. Control charts based on weighted sums. *The Annals of Statistics*, 2(1):134–147, 1974.
- [60] T. Lai. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(4):613–644, 1995.

- [61] T. Lai. Information bounds and quick detection of parameter changes in stochastic systems. *IEEE Transactions on Information Theory*, 44(7):2917–2929, 1998.
- [62] H. Landau and L. Shepp. On the supremum of a Gaussian process. *Sankhyā: The Indian Journal of Statistics, Series A*, 32:369–378, 1970.
- [63] M. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*, volume 21. Springer-Verlag New York, 1983.
- [64] B. Levin and J. Kline. The CUSUM test of homogeneity with an application in spontaneous abortion epidemiology. *Statistics in Medicine*, 4(4):469–488, 1985.
- [65] S. Li. Concise formulas for the area and volume of a hyperspherical cap. *Asian Journal of Mathematics and Statistics*, 4(1):66–70, 2011.
- [66] W. Li and Q. Shao. Lower tail probabilities for Gaussian processes. *The Annals of Probability*, 32(1):216–242, 2004.
- [67] J. Lin. Approximating the normal tail probability and its inverse for use on a pocket calculator. *Applied Statistics*, 38(1):69–70, 1989.
- [68] G. Lorden. Procedures for reacting to a change in distribution. *The Annals of Mathematical Statistics*, 42(6):1897–1908, 1971.
- [69] M. Marcus and L. Shepp. Sample behavior of Gaussian processes. In *Proc. of the Sixth Berkeley Symposium on Math. Statist. and Prob.*, 2:423–421, 1972.
- [70] S. McAleese. *Operational aspects of oil and gas well testing*. Elsevier, 2000.
- [71] C.B. Mehr and J.A. McFadden. Certain properties of Gaussian processes and their first-passage times. *Journal of the Royal Statistical Society. Series B (Methodological)*, 27(3):505–522, 1965.
- [72] J. Mohamed and L. Delves. *Computational Methods for Integral Equations*. Cambridge University Press, 1985.
- [73] G. Molchan. Survival exponents for some Gaussian processes. *International Journal of Stochastic Analysis*, 2012.
- [74] V. Moskvina and A. Zhigljavsky. An algorithm based on Singular Spectrum Analysis for change-point detection. *Communications in Statistics—Simulation and Computation*, 32(2):319–352, 2003.
- [75] G. Moustakides. Optimal stopping times for detecting changes in distributions. *Annals of Statistics*, 14(4):1379–1387, 1986.
- [76] G. Moustakides. Optimality of the CUSUM procedure in continuous time. *The Annals of Statistics*, 32(1):302–315, 2004.

- [77] G. Moustakides, A. Polunchenko, and A. Tartakovsky. Numerical comparison of CUSUM and Shiryaev–Roberts procedures for detecting changes in distributions. *Communications in Statistics—Theory and Methods*, 38(16-17): 3225–3239, 2009.
- [78] J. Naus. Approximations for distributions of scan statistics. *Journal of the American Statistical Association*, 77(377):177–183, 1982.
- [79] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*. SIAM, 1992.
- [80] J. Noonan and A. Zhigljavsky. Approximations of the boundary crossing probabilities for the maximum of moving weighted sums. *Statistical Papers*, 59(4): 1325–1337, 2018.
- [81] J. Noonan and A. Zhigljavsky. Approximating Shepp’s constants for the Slepian process. *Statistics & Probability Letters*, 2019.
- [82] J. Noonan and A. Zhigljavsky. Approximations for the boundary crossing probabilities of moving sums of normal random variables. *Communications in Statistics-Simulation and Computation*, pages 1–22, 2019.
- [83] J. Noonan and A. Zhigljavsky. Approximations for the Boundary Crossing Probabilities of Moving Sums of Random Variables. *Methodology and Computing in Applied Probability*, 2020.
- [84] J. Noonan and A. Zhigljavsky. Power of the MOSUM test for online detection of a transient change in mean. *Sequential Analysis*, 39(2):269–293, 2020.
- [85] J. Noonan and A. Zhigljavsky. Efficient quantization and weak covering of high dimensional cubes. *arXiv preprint arXiv:2005.07938*, 2020.
- [86] J. Noonan and A. Zhigljavsky. Non-lattice covering and quantization in high dimensions. In *Black Box Optimization, Machine Learning and No-Free Lunch Theorems*. Springer, 2021.
- [87] E. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.
- [88] V. Petrov. *Sums of independent random variables*. Springer-Verlag, 1975.
- [89] V. Petrov. *Limit theorems of probability theory: sequences of independent random variables*. Oxford Science Publications, 1995.
- [90] J. Pickands. Upcrossing probabilities for stationary Gaussian processes. *Transactions of the American Mathematical Society*, 145:51–73, 1969.
- [91] J. Pitman and W. Tang. The Slepian zero set, and Brownian bridge embedded in Brownian motion by a spacetime shift. *Electronic Journal of Probability*, 20(61):1–28, 2015.
- [92] M. Pollak. Optimal detection of a change in distribution. *The Annals of Statistics*, 13(1):206–227, 1985.

- [93] M. Pollak. Average run lengths of an optimal method of detecting a change in distribution. *The Annals of Statistics*, 15(2):749–779, 1987.
- [94] M. Pollak and D. Siegmund. A diffusion process and its applications to detecting a change in the drift of Brownian motion. *Biometrika*, 72(2):267–280, 1985.
- [95] M. Pollak and D. Siegmund. Convergence of quasi-stationary to stationary distributions for stochastically monotone Markov processes. *Journal of Applied Probability*, 23(1):215–220, 1986.
- [96] M. Pollak and A. Tartakovsky. Optimality properties of the Shiryaev–Roberts procedure. *Statistica Sinica*, 19(4):1729–1739, 2009.
- [97] A. Polunchenko. On the quasi-stationary distribution of the Shiryaev–Roberts diffusion. *Sequential Analysis*, 36(1):126–149, 2017.
- [98] A. Polunchenko. Asymptotic near-minimaxity of the randomized Shiryaev–Roberts–Pollak change-point detection procedure in continuous time. *Theory of Probability & Its Applications*, 62(4):617–631, 2018.
- [99] A. Polunchenko and A. Tartakovsky. On optimality of the Shiryaev–Roberts procedure for detecting a change in distribution. *The Annals of Statistics*, 38(6):3445–3457, 2010.
- [100] A. Polunchenko, A. Tartakovsky, and N. Mukhopadhyay. Nearly optimal change-point detection with an application to cybersecurity. *Sequential Analysis*, 31(3):409–435, 2012.
- [101] V. Poor. *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.
- [102] V. Poor and O. Hadjiladis. *Quickest detection*. Cambridge University Press, 2008.
- [103] B. Prakasa Rao. *Asymptotic theory of statistical inference*. Wiley, 1987.
- [104] M. Priestley. *Spectral analysis and time series*. Academic press, 1981.
- [105] L. Pronzato and W. G. Müller. Design of computer experiments: space filling and beyond. *Statistics and Computing*, 22(3):681–701, 2012.
- [106] A. Račkauskas and C. Suquet. Hölder norm test statistics for epidemic change. *Journal of statistical planning and inference*, 126(2):495–520, 2004.
- [107] M. Reed and B. Simon. *Methods of Modern Mathematical Physics: Scattering theory Vol. 3*. Academic Press, 1979.
- [108] Y. Ritov. Decision theoretic optimality of the CUSUM procedure. *The Annals of Statistics*, 18(3):1464–1469, 1990.

- [109] S. Roberts. A comparison of some control chart procedures. *Technometrics*, 8(3):411–430, 1966.
- [110] S. Roberts. Control chart tests based on geometric moving averages. *Technometrics*, 42(1):97–101, 2000.
- [111] L. Schmerr. *Fundamentals of ultrasonic nondestructive evaluation*. Springer, 2016.
- [112] L. Shepp. Radon-Nikodym derivatives of Gaussian measures. *The Annals of Mathematical Statistics*, 37(2):321–354, 1966.
- [113] L. Shepp. First passage time for a particular Gaussian process. *The Annals of Mathematical Statistics*, 42(3):946–951, 1971.
- [114] L. Shepp and D. Slepian. First-passage time for a particular stationary periodic Gaussian process. *Journal of Applied Probability*, 13(1):27–38, 1976.
- [115] W. Shewhart. *Economic control of quality of manufactured product*. Macmillan And Co Ltd, London, 1931.
- [116] A. Shiryaev. On optimum methods in quickest detection problems. *Theory of Probability & Its Applications*, 8(1):22–46, 1963.
- [117] D. Siegmund. *Sequential Analysis: Tests and Confidence Intervals*. Springer Science & Business Media, 1985.
- [118] D. Siegmund. Boundary crossing probabilities and statistical applications. *The Annals of Statistics*, 14(2):361–404, 1986.
- [119] D. Siegmund. Approximate tail probabilities for the maxima of some random fields. *The Annals of Probability*, 16(2):487–501, 1988.
- [120] D. Siegmund and E. Venkatraman. Using the generalized likelihood ratio statistic for sequential detection of a change-point. *The Annals of Statistics*, 23(1):255–271, 1995.
- [121] D. Siegmund, N. Zhang, and B. Yakir. False discovery rate for scanning statistics. *Biometrika*, 98(4):979–985, 2011.
- [122] D. Slepian. First passage time for a particular Gaussian process. *The Annals of Mathematical Statistics*, 32(2):610–612, 1961.
- [123] D. Slepian. The one-sided barrier problem for Gaussian noise. *Bell System Technical Journal*, 41(2):463–501, 1962.
- [124] I. Sobol'. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational mathematics and mathematical physics*, 7(4):784–802, 1967.
- [125] I. Sobol'. Uniformly distributed sequences with additional uniformity properties. *USSR Computational mathematics and mathematical physics*, 16(5):1332–1337, 1976.

- [126] R. Streit and P. Willett. Detection of random transient signals via hyperparameter estimation. *IEEE transactions on signal processing*, 47(7):1823–1834, 1999.
- [127] A. Sukharev. Optimal strategies of the search for an extremum. *USSR Computational Mathematics and Mathematical Physics*, 11(4):119–137, 1971.
- [128] A. Sukharev. *Minimax models in the theory of numerical methods*. Springer Science & Business Media, 1992.
- [129] A. Tartakovsky. Asymptotic performance of a multichart CUSUM test under false alarm probability constraint. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 320–325. IEEE, 2005.
- [130] A. Tartakovsky. Discussion on “is average run length to false alarm always an informative criterion?” by Yajun Mei. *Sequential Analysis*, 27(4):396–405, 2008.
- [131] A. Tartakovsky, I. Nikiforov, and M. Basseville. *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.
- [132] B. Tibken, D. Constales, and et al. The volume of the intersection of a concentric cube and ball in n-dimensional space: collection of approximations. In *SIAM Review*, volume 39, pages 783–786, 1997.
- [133] G. Tóth. Packing and covering. In *Handbook of discrete and computational geometry*, pages 27–66. Chapman and Hall/CRC, 2017.
- [134] G. Tóth and W. Kuperberg. Packing and covering with convex sets. In *Handbook of Convex Geometry*, pages 799–860. Elsevier, 1993.
- [135] M. Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.
- [136] K. Waldmann. Bounds to the distribution of the run length in general quality-control schemes. *Statistische Hefte*, 27(1):37, 1986.
- [137] X. Wang and J. Glaz. Variable window scan statistics for normal data. *Communications in Statistics-Theory and Methods*, 43(10-12):2489–2504, 2014.
- [138] X. Wang, B. Zhao, and J. Glaz. A multiple window scan statistic for time series models. *Statistics & Probability Letters*, 94:196–203, 2014.
- [139] Q. Yao. Large deviations for boundary crossing probabilities of some random fields. In *Journal of Mathematical Research and Exposition*, volume 9, pages 181–192, 1989.
- [140] Q. Yao. Boundary-crossing probabilities of some random fields related to likelihood ratio tests for epidemic alternatives. *Journal of Applied Probability*, 30(1):52–65, 1993.

- [141] Q. Yao. Tests for change-points with epidemic alternatives. *Biometrika*, 80(1): 179–191, 1993.
- [142] A. Zhigljavsky. *Theory of global random search*. Kluwer Academic Publishers, 1991.
- [143] A. Zhigljavsky and A. Kraskovsky. *Detection of abrupt changes of random processes in radiotechnics problems*. St. Petersburg University Press, 1988. (in Russian).
- [144] A. Zhigljavsky and J. Noonan. Covering of high-dimensional cubes and quantization. In *SN Operations Research Forum*, volume 1, pages 1–32. Springer, 2020.
- [145] A. Zhigljavsky and J. Noonan. First passage times for Slepian process with linear and piecewise linear barriers. *Extremes*, pages 1–25, 2021.
- [146] A. Zhigljavsky and A. Žilinskas. *Stochastic global optimization*. Springer Science & Business Media, 2007.
- [147] A. Žilinskas. On the worst-case optimal multi-objective global optimization. *Optimization Letters*, 7(8):1921–1928, 2013.